Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 9(8): 23088–23099.
DOI: 10.3934/math.20241122
Received: 22 February 2024
Revised: 08 July 2024
Accepted: 22 July 2024
Published: 29 July 2024

Research article

Weighted expectile average estimation based on CBPS with responses missing at random

Qiang Zhao[1], **Zhaodi Wang**[1], **Jingjing Wu**[2] **and Xiuli Wang**[1,*]

[1] School of Mathematics and Statistics, Shandong Normal University, Jinan 250014, China
[2] Department of Mathematics and Statistics, University of Calgary, Calgary, AB, Canada

* **Correspondence:** Email: wxlmath@163.com.

Abstract: An improved weighted expectile average estimator for the regression coefficient has been obtained based on the covariate balancing propensity score (CBPS), when the responses of linear models are missing at random. The asymptotic normality of the proposed method has been proved, and the estimation effect of the method is further illustrated by numerical simulation.

Keywords: expectile regression; weighted expectile average estimator; covariate balancing propensity score; responses missing; generalized method of moments
Mathematics Subject Classification: 62F10, 62F12

1. Introduction

In practical applications, due to the interference of various factors, collected data is often incomplete. Missing data is common in public opinion polls, medical research, experimental science, and other application fields. Missing data will not only result in the reduction of effective information, the deviation of the estimation result, but also affect the statistical decision-making and distort the analysis result to some extent. One approach to deal with missing data is complete-case analysis, which deletes all incomplete data. However, Little and Rubin [1] pointed out that this will cause biased estimation when the occurrence of missing data is not completely at random. Yates [2] introduced an imputation method which is widely used to handle missing responses. The purpose of this method is to find suitable values for the missing data to impute. Then, the data of the filled values are regarded as the complete observation data, which can be analyzed by the classical method. Inverse probability weighting (IPW), which was proposed by Horvitz and Thompson [3], is another method to deal with missing data. The inverse of the selection probability is chosen to be the weight assigned to the fully observed data. The missing at random (MAR) assumption, in the sense of Rubin et al. [4], is a common assumption for statistical analysis with missing data.

In the case of missing data, the missing mechanism is usually unknown, and parameter methods and nonparametric methods are commonly used to estimate. For the parameter method, there may be a model misspecification problem. Imai and Ratkovic [5] proposed the covariate balanced propensity score (CBPS), which improves the parameter method. Based on the CBPS method, Guo et al. [6] applied the CBPS method to mean regression to obtain the estimators of the regression parameters β and the mean μ in the case of missing data.

Expectile regression, which was proposed by Newey and Powell [7], can be regarded as a generalization of mean regression. Expectile regression uses the sum of asymmetric residual squares as the loss function, and since the loss function is convex and differentiable, expectile regression has computational advantages over quantile regression. Recently, people have carried out a lot of specific research on expectile regression. Sobotka et al. [8] established the asymptotic properties of a semi-parametric expectile regression estimator and introduced confidence intervals for expectiles. Waltrup et al. [9] observed that expectile regression tends to have less crossing and more robustness against heavy tailed distributions than quantile regression. Ziegel [10] concluded that expectile shares coherence and elicitability. Pan et al. [11] considered fitting a linear expectile regression model for estimating conditional expectiles based on a large quantity of data with covariates missing at random. Recently, Pan et al. [12] developed a weighted expectile regression approach for estimating the conditional expectile when covariates are missing at random (MAR). They only considered a single expectile, and the missing mechanism was assumed to be logistic regression. However, the missing mechanism model may be misspecified. In addition, it is known that making full use of multiple target information can improve the efficiency of parameter estimation. In summary, when the model may be misspecified, we use the idea of covariate balance to study the weighted expectile average estimation of unknown parameters based on CBPS by using multiple expected information. Our estimators can improve performance of the usual weighted expectile average estimator in terms of standard deviation (SD) and mean squared error (MSE).

The rest of this paper is organized as follows. In Section 2, we propose a CBPS-based estimator for the propensity score. In Section 3, we estimate the expected quantile weighted average of the regression parameters based on CBPS. Moreover, we establish the asymptotic normality of the weighted estimator in Section 4. In Section 5, a simulation study is carried out to assess the performance of the proposed method. The proofs of those theoretical results are deferred to the Appendix.

2. CBPS-based estimator for the propensity score

Consider the following linear regression model:

$$Y_i = X_i^T\beta + \varepsilon_i, \quad i = 1, 2, \dots, n, \tag{2.1}$$

where Y_i is response, X_i is covariate, β is the p-dimensional vector of unknown parameters, and ε_i is the random error. Assuming that the response variable Y_i is missing at random, the covariate X_i can be fully observed. For the i^{th} individual, let δ_i denote the observing indicator, i.e., $\delta_i = 1$ if Y_i is observed and 0 otherwise. In our paper, we only consider the missing mechanism of missing at random (MAR), that is,

$$P(\delta_i = 1|X_i, Y_i) = \pi(X_i) \triangleq \pi_i, \tag{2.2}$$

where π_i is called the selection probability function or the propensity score.

The most popular choice of $\pi(X_i)$ is a logistic regression function (Peng et al. [13]). We make the same choice and posit a logistic regression model for $\pi(X_i)$,

$$\pi(X_i,\gamma)=\frac{\exp(\gamma_0+X_i^T\gamma_1)}{1+\exp(\gamma_0+X_i^T\gamma_1)},\tag{2.3}$$

and $\gamma=(\gamma_0,\gamma_1^T)^T\in\Theta$ is the unknown parameter vector with the parameter space $\Theta\subseteq R^{q+1}$. Here, γ can be estimated by maximizing the log-likelihood function

$$L(\gamma)=\sum_{i=1}^{n}\{\delta_i\log\pi(X_i,\gamma)+(1-\delta_i)\log(1-\pi(X_i,\gamma)\}.$$

Assuming that $\pi(X_i,\gamma)$ is twice continuously differentiable with respect to γ, maximizing $L(\gamma)$ implies the first-order condition

$$\frac{1}{n}\sum_{i=1}^{n}s(\delta_i,X_i,\gamma)=0,\quad s(\delta_i,X_i,\gamma)=\frac{\delta_i\pi'(X_i,\gamma)}{\pi(X_i,\gamma)}-\frac{(1-\delta_i)\pi'(X_i,\gamma)}{1-\pi(X_i,\gamma)},\tag{2.4}$$

where $\pi'(X_i,\gamma)=\partial\pi(X_i,\gamma)/\partial\gamma^T$. The maximum likelihood method is a commonly used and simple parameter estimation method. However, when the selection probability model (2.3) is assumed to be wrong, the estimator based on this method will have a large deviation. In order to make the parameter method more robust, we use the covariate balanced propensity score method proposed by Imai and Ratkovic [5] to estimate the unknown parameter γ, that is,

$$E\left\{\frac{\delta_i\tilde{X}_i}{\pi(X_i,\gamma)}-\frac{(1-\delta_i)\tilde{X}_i}{1-\pi(X_i,\gamma)}\right\}=0.\tag{2.5}$$

$\tilde{X}_i=f(X_i)$ is an M-dimensional vector-valued measurable function of X_i. For any covariate function, as long as the expectation exists, Eq (2.5) must hold. If the propensity score model is incorrectly specified, then the maximum likelihood may not be able to balance the covariates. Following Imai and Ratkovic [5], we can set $\tilde{X}_i=X_i$ to ensure that the first moment of each covariate is balanced even when the model is misspecified. $\pi(X_i,\gamma)$ satisfies the condition

$$E\left\{\frac{\delta_iX_i}{\pi(X_i,\gamma)}-\frac{(1-\delta_i)X_i}{1-\pi(X_i,\gamma)}\right\}=0.\tag{2.6}$$

The sample form of the covariate equilibrium condition obtained from (2.6) is

$$\frac{1}{n}\sum_{i=1}^{n}z(\delta_i,X_i,\gamma)X_i=0,\tag{2.7}$$

where

$$z(\delta_i,X_i,\gamma)=\frac{\delta_i-\pi(X_i,\gamma)}{\pi(X_i,\gamma)(1-\pi(X_i,\gamma))}.$$

According to Imai and Ratkovic [5], if we only use the condition of the $\pi'(X_i,\gamma)$ equilibrium, i.e., (2.4), at this time, the number of equations is equal to the number of parameters. Then, the covariate equilibrium propensity score is just-identified. If we combine Eq (2.4) with the score condition given in Eq (2.7),

$$\bar{U}(\gamma) = \frac{1}{n}\sum_{i=1}^{n} U(\delta_i, X_i, \gamma), \tag{2.8}$$

where

$$U(\delta_i, X_i, \gamma) = \begin{pmatrix} s(\delta_i, X_i, \gamma) \\ z(\delta_i, X_i, \gamma)X_i \end{pmatrix},$$

then the covariate equilibrium propensity score is over-identified because the number of moment conditions exceeds that of parameters. For over-identified CBPS, the estimation of γ can be obtained by using the generalized moment method (GMM) (Hansen [14]). For a positive semidefinite symmetric weight matrix W, the GMM estimator $\hat{\gamma}$ can be obtained by minimizing the following objective function for γ:

$$Q(\gamma) = \bar{U}^T(\gamma) W \bar{U}(\gamma). \tag{2.9}$$

The above method is also applicable to the case where the covariate balanced propensity score is just-identified.

3. Estimator for the regression parameter

Pan et al. [12] introduced the weighted expectile regression estimation of a linear model in detail. According to the idea of inverse probability weighting, when the selection probability function $(\pi_1 \ldots, \pi_n)^T$ is known, the expectile estimator of β under missing responses is defined as

$$(\hat{\beta}_{\tau_k,T}, \hat{b}_{\tau_k}) = \underset{\beta, b_{\tau_k}}{\arg\min} \sum_{i=1}^{n} \frac{\delta_i}{\pi(X_i, \gamma)} \Phi_{\tau_k}(Y_i - X_i^T\beta - b_{\tau_k}), \tag{3.1}$$

where $\tau_k \in (0, 1)$ is expectile level, and $\Phi_{\tau_k}(v) = |\tau_k - I(v \leq 0)|\, v^2$. b_{τ_k} represents the τ_k-expectile of the error term ε_i. Then, according to Zhao et al. [15], let K be the number of expectiles, and consider the equally spaced expectiles $\tau_k = \frac{k}{K+1}$, $k = 1, 2, \ldots, K$. The weighted expectile average estimator of the linear model parameter β when the missing mechanism is known is defined as

$$\hat{\beta} = \sum_{k=1}^{K} \omega_k \hat{\beta}_{\tau_{k,T}},$$

where the weight vector $(\omega_1, \ldots, \omega_K)^T$ satisfies $\sum_{k=1}^{K} \omega_k = 1$.

When the selection probability function is unknown, we use the method proposed in the second section to estimate the parameter γ based on CBPS, so as to obtain $\pi(X_i, \hat{\gamma})$. The loss function of the τ_k-expectile can be defined as

$$L_n(\beta_{\tau_k}, b_{\tau_k}) = \sum_{i=1}^{n} \frac{\delta_i}{\pi(X_i, \hat{\gamma})} \Phi_{\tau_k}(Y_i - X_i^T\beta - b_{\tau_k}).$$

By minimizing the loss function, we can obtain the expectile estimation of the unknown parameter β,

$$(\hat{\beta}_{\tau_k}, \hat{b}_{\tau_k}) = \underset{\beta, b_{\tau_k}}{\arg\min}\, L_n(\beta_{\tau_k}, b_{\tau_k}). \tag{3.2}$$

Therefore, the weighted expectile average estimation of the linear model parameter β when the missing mechanism is unknown under the missing responses is defined as

$$\hat{\beta}_w = \sum_{k=1}^{K} \omega_k \hat{\beta}_{\tau_k}. \tag{3.3}$$

The weight vector $(\omega_1, \ldots, \omega_K)^T$ satisfies $\sum_{k=1}^{K} \omega_k = 1$.

4. Asymptotic property

Let γ_0 and β_0 represent the true values of γ and β respectively, and $U(\gamma) = \begin{pmatrix} s(\delta, X, \gamma) \\ z(\delta, X, \gamma)X \end{pmatrix}$. In addition, with reference to Pan et al. [12] and Guo [16], the following regularity conditions are required.

C1: γ_0 is the interior point of Θ.

C2: $U(\gamma)$ is differentiable in the neighborhood $\triangle$ of γ_0.

C3: $E[U(\gamma_0)] = 0$, $E[||U(\gamma_0)||^2] < \infty$.

C4: $E[\sup_{\gamma\in\triangle} \left\|\nabla_\gamma U(\gamma)\right\|] < \infty$, where ∇_γ is the first-order partial derivative of the function to γ.

C5: $\Gamma = E[\nabla_\gamma U(\gamma)]$ exists.

C6: For any i, there exists a compact set $\mathcal{X}$, such that $X_i \in \mathcal{X} \subset \mathbb{R}^p$, and X_i and ε_i are independent.

C7: The regression errors $\{\varepsilon_i\}_{i=1}^n$ are independent and identically distributed with common cumulative distribution function $F(\cdot)$, satisfying $E[\varepsilon_i^2] < \infty$.

C8: There exists $a > 0$ such that $\pi(V_i, \gamma) > a$ for any i.

C9: The symmetric matrix Σ_1 is positive definite.

The following theorem presents the asymptotic distribution of $\hat{\beta}_w$.

Theorem 4.1 (Asymptotic Normality of $\hat{\beta}_w$) Under the assumptions C1–C9, we have

$$\sqrt{n}(\hat{\beta}_\omega - \beta_0) \xrightarrow{d} N(0, \Sigma_1^{-1}\Lambda\Sigma_1^{-1}),$$

where $\Sigma_1 = E[X_i X_i^T]$, $\Lambda = E[\lambda\lambda^T], \lambda = \mu - E[\partial\mu/\partial\gamma^T]\{E[\partial U(\gamma)/\partial\gamma^T]\}^{-1}U(\gamma)$, $\mu = \frac{\delta}{\pi(X,\gamma)} X \sum_{k=1}^{K} \frac{\omega_k \Psi_{\tau_k}(\varepsilon - b_{0k})}{g(\tau_k)}$.

5. Simulation

In the following, the expectile weighted average estimation based on covariate balancing propensity score proposed in this paper is analyzed by numerical simulation, and the method is compared with the usual parameter estimation method in the case of correct and wrong model assumptions. Consider the following linear model:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon, \tag{5.1}$$

where $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1$, and (X_1, X_2, X_3) obeys the joint normal distribution with mean of 0, covariance of 0.5, and variance of 1. The error term ε obeys the standard normal distribution. In our simulation, we take $K = 10$, $\tau_k = k/11$ for $k = 1, 2, \ldots, 10$, and consider the real choice probability model as

$$\pi(X_1, X_2, X_3) = \exp(0.3X_1 + 0.25X_2 + 0.25X_3)/[1 + \exp(0.3X_1 + 0.25X_2 + 0.25X_3)]. \tag{5.2}$$

Under the assumption of random missing, in order to illustrate the effect when the model is misspecified, we assume that the covariates

$$X^* = (X_1^*, X_2^*, X_3^*) \triangleq \{\exp(X_1/2), (X_2)/\{1 + \exp(X_1)\} + 10, (X_1 X_3/25 + 0.6)^3\}.$$

If the model (5.2) is represented by $\pi(X^*)$, the model will be specified incorrectly. In the simulation study of the expectile regression of the unknown parameter β, we consider the following two cases: (1) Propensity score model is correctly specified. (2) Propensity score model is misspecified. Zhao [15] proposed the weighted composite expectile regression method for a varying-coefficient partially linear model. For a given scenario, referring to Zhao [15], we compare the weighted expectile average estimation based on CBPS, denoted as CBPS-WEAE, with weighted composite expectile regression, denoted as WCER, and weighted composite quantile regression, denoted as WCQR, to examine the performance of the estimator, where the weights of WCER and WCQR are estimated by the generalized linear model.

In the simulation, samples of size $n = 500, 800, 1000, 1200$ are generated independently. For each scenario, we conduct 1000 simulations and calculate the average mean squared error (MSE) for estimator of β and the average bias (Bias) and standard deviation (SD) for estimator of β_1, β_2, and β_3. In order to examine the influence of the error distribution on the performance of the proposed method, two different distributions of the model error ε are considered: standard normal distribution $N(0, 1)$ and centralized χ^2 distribution with 4 degrees of freedom. The results of our simulations are presented in Tables 1 and 2.

From Tables 1 and 2 we observe that, as expected, all three estimators are unbiased. In terms of MSE, as a convenient measure of average error, we observe that when model error ε follows the standard normal distribution $N(0, 1)$, CBPS-WEAE performs best among the three estimators considered, followed immediately by WCER, while WCQR performs worst. When ε follows a centralized χ^2 distribution with 4 degrees of freedom, CBPS-WEAE is superior to the other two methods. When sample size is large, it can be seen that the performance of the three estimators is significantly improved compared with that when the sample size is small. In general, our proposed improved estimator is effective.

Table 1. Simulation results (×100) under the error $\varepsilon \sim N(0, 1)$.

n	Model	Method	MSE	β_1		β_2		β_3	
				Bias	SD	Bias	SD	Bias	SD
500	correct	WCQR	2.437	-0.100	9.038	-0.221	9.041	2.498	8.966
		WCER	2.155	-0.172	8.533	-0.303	8.384	0.366	8.512
		CBPS-WEAE	2.139	0.023	8.490	-0.433	8.355	-0.218	8.488
	incorrect	WCQR	2.371	0.731	8.908	-0.616	8.866	2.498	8.855
		WCER	2.256	0.518	8.642	-0.471	8.698	0.366	8.658
		CBPS-WEAE	2.122	0.348	8.382	-0.680	8.547	-0.104	8.280
800	correct	WCQR	1.490	-0.033	6.944	0.105	7.190	2.498	7.011
		WCER	1.380	-0.012	6.616	-0.036	6.886	0.366	6.844
		CBPS-WEAE	1.356	0.311	6.569	-0.219	6.931	0.076	6.663
	incorrect	WCQR	1.392	0.474	6.729	-0.176	6.689	2.498	6.997
		WCER	1.357	0.266	6.980	0.291	6.455	0.366	6.732
		CBPS-WEAE	1.310	-0.262	6.676	0.098	6.536	-0.285	6.609
1000	correct	WCQR	1.491	0.123	6.156	-0.143	6.427	2.498	6.375
		WCER	1.107	0.067	6.008	0.003	6.182	0.366	6.044
		CBPS-WEAE	1.098	3.303	6.094	-0.296	5.973	-0.260	6.069
	incorrect	WCQR	1.202	0.037	6.497	-0.307	6.477	2.498	6.000
		WCER	1.172	-0.155	6.252	0.213	6.452	0.366	6.042
		CBPS-WEAE	1.122	-0.483	6.137	-0.070	6.209	-0.279	5.985
1200	correct	WCQR	1.033	0.021	5.819	0.136	5.967	2.498	5.824
		WCER	0.902	0.132	5.513	0.027	5.470	0.366	5.472
		CBPS-WEAE	0.898	0.403	5.486	-0.115	5.487	-0.267	5.421
	incorrect	WCQR	1.005	0.401	5.968	-0.347	5.687	2.498	5.681
		WCER	0.960	0.117	5.682	-0.027	5.578	0.366	5.712
		CBPS-WEAE	0.923	-0.118	5.611	-0.290	5.451	-0.119	5.571

Table 2. Simulation results (×100) under the error $\varepsilon \sim \chi^2(4)$.

n	Model	Method	MSE	β_1		β_2		β_3	
				Bias	SD	Bias	SD	Bias	SD
500	correct	WCQR	42.007	1.385	36.000	1.970	37.445	2.498	38.742
		WCER	19.889	0.282	25.231	0.767	26.285	0.366	25.743
		CBPS-WEAE	18.424	-3.970	24.657	10.138	24.950	-0.236	24.431
	incorrect	WCQR	37.935	4.173	36.241	0.443	35.828	2.498	34.384
		WCER	19.472	1.481	25.942	-0.235	25.159	0.366	25.316
		CBPS-WEAE	19.078	-3.454	25.124	-3.294	24.263	-4.148	25.491
800	correct	WCQR	33.040	1.769	33.706	1.452	32.203	2.498	33.593
		WCER	12.696	1.142	21.067	0.887	20.695	0.366	19.901
		CBPS-WEAE	12.471	-4.204	21.122	1.985	19.963	-0.534	19.540
	incorrect	WCQR	32.691	2.495	34.420	1.215	32.328	2.498	32.158
		WCER	13.098	-0.743	21.348	0.776	20.046	0.366	21.269
		CBPS-WEAE	12.594	-4.931	20.047	-2.309	20.593	-3.187	19.872
1000	correct	WCQR	31.334	2.961	32.292	-1.051	32.338	2.498	32.208
		WCER	12.647	1.554	21.038	0.214	19.026	0.366	21.370
		CBPS-WEAE	9.456	-3.280	18.555	0.380	16.798	0.746	17.568
	incorrect	WCQR	31.671	4.760	33.762	0.512	31.422	2.498	31.931
		WCER	11.049	0.102	18.822	-0.729	17.908	0.366	20.694
		CBPS-WEAE	9.811	-2.869	18.456	-3.676	17.077	-2.495	17.939
1200	correct	WCQR	29.885	-0.103	31.443	2.665	30.750	2.498	32.389
		WCER	11.241	-0.516	18.673	1.694	18.976	0.366	20.328
		CBPS-WEAE	8.751	-5.023	17.754	1.652	16.251	-0.198	16.391
	incorrect	WCQR	31.091	1.617	33.886	3.476	30.736	2.498	31.688
		WCER	10.297	0.207	18.428	0.508	18.222	0.366	18.942
		CBPS-WEAE	9.455	-4.175	17.484	-2.462	16.866	-3.193	17.961

6. Conclusions

In this paper, in order to improve the estimation efficiency of weighted expectile average estimation, we estimate the selection probability function based on CBPS and propose a weighted expectile average estimator based on CBPS when the response variables are missing at random. The asymptotic normality of the proposed method is proved, and the estimation effect of the method is further illustrated by numerical simulation. The numerical simulation results show that the method is effective.

Author contributions

Qiang Zhao: Conceptualization, methodology, supervision, writing-review and editing; Zhaodi Wang: Validation, software, writing-original draft; Jingjing Wu: Funding acquisition, formal analysis, writing-original draft; Xiuli Wang: Funding acquisition, investigation, resources, writing-review and editing. All authors have read and approved the final version of the manuscript for publication.

Use of AI tools declaration

The authors declare they have not used artificial intelligence (AI) tools in the creation of this article.

Acknowledgments

The research is supported by Natural Science Foundation of Shandong Province (Grant Nos. ZR2021MA077 and ZR2021MA048).

Conflict of interest

All authors declare that there is no conflict of interest.

References

1. R. J. A. Little, D. B. Rubin, *Statistical analysis with missing data*, 2 Eds., New York: Wiley, 2002. http://dx.doi.org/10.1002/9781119013563
2. F. Yates, The analysis of replicated experiments when the field results are incomplete, *Emprie Jour. Exp. Agric.*, **1** (1933), 129–142.
3. D. G. Horvitz, D. J. Thompson, A generalization of sampling without replacement from a finite universe, *J. Am. Stat. Assoc.*, **47** (1952), 663–685. http://dx.doi.org/10.1080/01621459.1952.10483446
4. J. M. Robins, A. Rotnitzky, L. P. Zhao, Estimation of regression coefficients when some of regression coefficients estimation regressors are not always observed, *J. Am. Stat. Assoc.*, **89** (1994), 846–866. http://dx.doi.org/10.2307/2290910
5. K. Imai, M. Ratkovic, Covariate balancing propensity score, *J. R. Stat. Soc. B.*, **76** (2014), 243–263. http://dx.doi.org/10.1111/rssb.12027
6. D. Guo, L. Xue, Y. Hu, Covariate-balancing-propensity-score-based inference for linear models with missing responses, *Statist. Probab. Lett.*, **123** (2017), 139–145. http://dx.doi.org/10.1016/j.spl.2016.12.001
7. W. K. Newey, J. L. Powell, Asymmetric least squares estimation and testing, *Econometrica*, **55** (1987), 819–847. http://dx.doi.org/10.2307/1911031
8. F. Sobotka, G. Kauermann, L. S. Waltrup, T. Kneib, On confidence intervals for semiparametric expectile regression, *Stat. Comput.*, **23** (2013), 135–148. http://dx.doi.org/10.1007/s11222-011-9297-1
9. L. S. Waltrup, F. Sobotka, T. Kneib, Expectile and quantile regression David and Goliath, *Stat. Model.*, **15** (2015), 433–456. http://dx.doi.org/10.1177/1471082X14561155
10. J. F. Ziegel, Coherence and elicitability, *Math. Financ.*, **26** (2016), 901–918. http://dx.doi.org/10.1111/mafi.12080

11. Y. Pan, Z. Liu, W. Cai, Large-scale expectile regression with covariates missing at random, *IEEE. Access.*, **8** (2020), 36502–36513. http://dx.doi.org/10.1109/access.2020.2970741

12. Y. Pan, Z. Liu, W. Cai, Weighted expectile regression with covariates missing at random, *Commun. Stat.-Simul. C.*, **52** (2023), 1057–1076. http://dx.doi.org/10.1080/03610918.2021.1873371

13. J. C. Peng, L. K. Lee, M. G. Ingersoll, An introduction to logistic regression analysis and reporting, *J. Educ. Res.*, **96** (2002), 3–14. http://dx.doi.org/10.1080/00220670209598786

14. L. P. Hansen, Large sample properties of generalized method of moments estimators, *Econometrica*, **50** (1982), 1029–1054. http://dx.doi.org/10.2307/1912775

15. S. Zhao, Expected regression estimation of semiparametric model, *Shanxi Normal Univ.*, 2021.

16. D. Guo, Estimation methods and theories of several types of regression models under missing data, *Beijing Univ. Tech.*, 2017.

17. N. L. Hjort, D. Pollard, Asymptotics for minimisers of convex processes, *Arxiv Preprint*, 2011. https://doi.org/10.48550/arXiv.1107.3806

Appendix: Assumptions and proofs

Define the following symbols:
$\eta_i = \frac{\delta_i}{\pi(X_i,\gamma)} X_i \Psi_{\tau_k}(\varepsilon_i)$,
$\hat{\eta}_i = \frac{\delta_i}{\pi(X_i,\hat{\gamma})} X_i \Psi_{\tau_k}(\varepsilon_i)$,
$F_n = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{\eta}_i$,
$\varepsilon_i = Y_i - X_i^T \beta_0$,
$\omega = (\omega_1, \omega_2, ...\omega_n)^T$,
$\Sigma_1 = E[X_i X_i^T]$,
$\Psi_{\tau_k} = 2\,|\tau_k - I(v \leq 0)|\, v$,
$u = (u_1, u_2, ..., u_n)^T$,
$G_n(u) = \sum_{i=1}^{n} \frac{\delta_i}{\pi(X_i,\hat{\gamma})} \Big[\Psi_{\tau_k}(\varepsilon_i - \frac{X_i^T u}{\sqrt{n}}) - \Psi_{\tau_k}(\varepsilon_i) \Big]$.

Lemma 1. Assume that C1–C5 hold. Then, when $n \to \infty$,

$$\sqrt{n}(\hat{\gamma} - \gamma_0) \xrightarrow{d} N(0, (\Gamma^T \Sigma^{-1} \Gamma)^{-1}),$$

where $\Gamma = E[\nabla_\gamma U(\gamma)]$, $\Sigma = E[U(\gamma) U^T(\gamma)]$.

The proof of Lemma 1 can refer to Theorem 2.2.1 in Guo [16].

Lemma 2. If the conditions C1–C4 are satisfied, then

$$F_n \xrightarrow{d} N(0, \Omega),$$

where $\Omega = E[QQ^T]$, $Q = \eta - E[\partial\eta/\partial\gamma^T]\{E[\partial U(\gamma)/\partial\gamma^T]\}^{-1} U(\gamma)$.

Proof. By expanding $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\eta}_i$ at γ and the proof process of Lemma 1, we can get

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\eta}_i &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i + \left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\eta_i}{\partial\gamma}\right]_{\gamma^*}\sqrt{n}(\hat{\gamma}-\gamma)\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i - \left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\eta_i}{\partial\gamma}\right]_{\gamma^*}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial U_i(\gamma)}{\partial\gamma}\right]_{\gamma^*}^{-1}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}U_i(\gamma)\right]\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\eta_i - D_nB_n^{-1}U_i(\gamma)\right],\end{aligned} \tag{A.1}$$

where $D_n = \left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\eta_i}{\partial\gamma}\right]_{\gamma^*}$, $B_n = \left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial U_i(\gamma)}{\partial\gamma}\right]_{\gamma^*}$, and γ^* lies between γ and $\hat{\gamma}$.

According to the central limit theorem,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\eta_i - D_nB_n^{-1}U_i(\gamma)) \xrightarrow{d} N(0,\Omega),$$

where $\Omega = E[QQ^T], Q = \eta - E[\partial\eta/\partial\gamma^T]\{E[\partial U(\gamma)/\partial\gamma^T]\}^{-1}U(\gamma)$.

Therefore, Lemma 2 is proved.

Lemma 3. If the conditions C1–C4 are satisfied, then

$$\sqrt{n}(\hat{\beta}_{\tau_k}-\beta_0) \xrightarrow{d} N(0, \frac{1}{4g^2(\tau)}\Sigma_1^{-1}\Omega\Sigma_1^{-1}).$$

Proof. If the conditions C1–C4 are satisfied, it can be known from Pan et al. [12] that

$$G_n(u) = g(\tau)u^T\Sigma_1 u + F_n^T u + o_p(1), \tag{A.2}$$

where $g(\tau) = (1-\tau)F(0) + \tau(1-F(0))$.

Known by Hjort and Pollard [17], if

$$D_n(u) = \frac{1}{2}u^TAu + B^Tu + o_p(1),$$

where $D_n(u)$ is a convex objective function with minimum point $\hat{u}_n$, A is a symmetric and positive definite matrix, and B is a random variable, then

$$\hat{u}_n \xrightarrow{d} -A^{-1}B.$$

Therefore, if we define $\hat{u}_n = \sqrt{n}(\hat{\beta}_{\tau_k}-\beta_0)$, then $\hat{\beta}_{\tau_k} = \beta_0 + \frac{\hat{u}_n}{\sqrt{n}}$. By some simple calculations and (A.2), we have

$$\begin{aligned}\hat{u}_n &= \arg\min_u \sum_{i=1}^{n}\frac{\delta_i}{\pi(V_i,\hat{\gamma})}\left[\Psi_{\tau_k}(\varepsilon_i - \frac{X_i^Tu}{\sqrt{n}}) - \Psi_{\tau_k}(\varepsilon_i)\right] = \arg\min_u G_n(u)\\ &= \arg\min_u\left[g(\tau)u^T\Sigma_1u + F_n^Tu + o_p(1)\right].\end{aligned} \tag{A.3}$$

According to condition C4, Σ_1 is a symmetric positive definite matrix. Lemma 3 is proved by Lemma 1 and Slutsky's theorem.

Proof of Theorem 4.1. By Lemma 3 we know that

$$\sqrt{n}(\hat{\beta}_{\tau_k}-\beta_0)=\Sigma_1^{-1}\frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{\pi(X_i,\hat{\gamma})}X_i\Psi_{\tau_k}(\varepsilon_i-b_{0k})}{2g(\tau_k)}+o_p(1).$$

From $\hat{\beta}_w=\sum_{k=1}^{K}\omega_k\hat{\beta}_{\tau_k}$, $\sum_{k=1}^{K}\omega_k=1$, we can get

$$\begin{aligned}\sqrt{n}(\hat{\beta}_w-\beta_0) &= \sqrt{n}(\sum_{k=1}^{K}\omega_k\hat{\beta}_{\tau_k}-\beta_0)\\ &= \sqrt{n}\sum_{k=1}^{K}\omega_k(\hat{\beta}_{\tau_k}-\beta_0)\\ &= \frac{1}{\sqrt{n}}\Sigma_1^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi(X_i,\hat{\gamma})}X_i\left\{\sum_{k=1}^{K}\frac{\omega_k\Psi_{\tau_k}(\varepsilon_i-b_{0k})}{2g(\tau_k)}\right\}+o_p(1).\end{aligned}\tag{A.4}$$

According to the proof of Lemma 2, we can obtain that

$$\sum_{i=1}^{n}\hat{\eta}_i=\sum_{i=1}^{n}\frac{\delta_i}{\pi(X_i,\hat{\gamma})}X_i\Psi_{\tau_k}(\varepsilon_i)=\sum_{i=1}^{n}\left[\eta_i-D_nB_n^{-1}U_i(\gamma)\right].$$

Let $\mu_i=\frac{\delta_i}{\pi(X_i,\gamma)}X_i\sum_{k=1}^{K}\frac{\omega_k\Psi_{\tau_k}(\varepsilon_i-b_{0k})}{g(\tau_k)}$, $\hat{\mu}_i=\frac{\delta_i}{\pi(X_i,\hat{\gamma})}X_i\sum_{k=1}^{K}\frac{\omega_k\Psi_{\tau_k}(\varepsilon_i-b_{0k})}{g(\tau_k)}$, and then

$$\sum_{i=1}^{n}\hat{\mu}_i=\sum_{i=1}^{n}\left[\mu_i-H_nB_n^{-1}U_i(\gamma)\right],$$

where $H_n=\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\mu_i}{\partial\gamma}\right]_{\gamma^*}$. Therefore, Eq (A.4) is equivalent to

$$\sqrt{n}(\hat{\beta}_w-\beta_0)=\frac{1}{\sqrt{n}}\Sigma_1^{-1}\sum_{i=1}^{n}\{\left[\mu_i-H_nB_n^{-1}U_i(\gamma)\right]\}+o_p(1).\tag{A.5}$$

Therefore,

$$\sqrt{n}(\hat{\beta}_\omega-\beta_0)\xrightarrow{d}N(0,\Sigma_1^{-1}\Lambda\Sigma_1^{-1}),$$

where $\Lambda=E[\lambda\lambda^T]$, $\lambda=\mu-E[\partial\mu/\partial\gamma^T]\{E[\partial U(\gamma)/\partial\gamma^T]\}^{-1}U(\gamma)$, $\mu=\frac{\delta}{\pi(X,\gamma)}X\sum_{k=1}^{K}\frac{\omega_k\Psi_{\tau_k}(\varepsilon-b_{0k})}{g(\tau_k)}$.