

Supplementary Material for “Bayesian Spatiotemporal Modeling for the Inpatient Hospital Costs of Alcohol-related Disorders”

A Discussions on composite quantile regression

This section provides a detailed motivation for imposing the equality-of-slope condition in the composite quantile regression and summarizes the findings on CQR in the existing literature.

In general, the regression of the response on covariates is given by

$$y_i = m(\mathbf{x}_i) + \epsilon_i, \quad i = 1, \dots, n, \quad (\text{A.1})$$

where $m(\mathbf{x}_i)$ is the conditional mean of y_i given the vector of covariates \mathbf{x}_i , ϵ_i is the error term with mean zero, and n is the sample size. Then, the τ -th conditional quantile function of y_i for $i = 1, \dots, n$ and $\tau \in (0, 1)$ is given by

$$Q_\tau(y_i | \mathbf{x}_i) = Q_\tau(m(\mathbf{x}_i)) + Q_\tau(\epsilon_i) \quad (\text{A.2})$$

$$\stackrel{\text{LR}}{=} \mathbf{x}_i^\top \boldsymbol{\beta}_\tau + Q_\tau(\epsilon_i) \quad (\text{A.3})$$

$$\stackrel{\text{if } \epsilon_i \text{'s are independent of } \mathbf{x}_i}{=} \mathbf{x}_i^\top \boldsymbol{\beta} + Q_\tau(\epsilon_i), \quad (\text{A.4})$$

where $Q_{\tau_l}(y_i | \mathbf{x}_i) = \inf\{y_i : F(y_i | \mathbf{x}_i) \geq \tau_l\}$, $Q_\tau(\cdot)$ are the τ -th conditional quantile function, and the second equality considers the case of linear regression (LR) when $m(\mathbf{x}_i) = \mathbf{x}_i^\top \boldsymbol{\beta}$.

It is worth emphasizing that when the error terms are independent of the covariates for LR, the slope coefficients of all quantile regressions converge in probability to the same vector (Koenker & Bassett, 1982), which yields the equality in (A.4). Then, the conditional quantile functions are actually a family of parallel hyperplanes with unknown parameters being one vector of slopes and a set of distinct intercepts (Koenker & Bassett, 1982). In this case, the fixed slopes across quantiles coincide with the covariate effects on the conditional mean. These observations motivate the combination of multiple quantile regressions in LR under the equality-of-slopes condition to obtain the weighted composite quantile regression (WCQR) estimators for robust and efficient inference on the conditional mean function (Koenker, 1984), i.e.,

$$(\hat{\alpha}_\tau, \hat{\boldsymbol{\beta}}) = \arg \min_{\alpha_\tau, \boldsymbol{\beta}} \sum_{l=1}^L \sum_{i=1}^n w_l \rho_{\tau_l}(y_i - \alpha_{\tau_l} - \mathbf{x}_i^\top \boldsymbol{\beta}), \quad (\text{A.5})$$

where $0 < \tau_1 < \dots < \tau_L < 1$ are L quantile levels, $\alpha_{\tau_l} = Q_{\tau_l}(\epsilon)$, w_l is a quantile-specific weight, and $\rho_{\tau_l}(u) = u\{\tau_l - 1(u < 0)\}$ is the quantile-specific check function for $l = 1, \dots, L$. Zou & Yuan (2008) viewed composite quantile regression (CQR) as an efficient and robust alternative to the least-squares (LS) estimator. They found that compared with the LS method, CQR has a relative efficiency greater than 70% regardless of the error distribution and sometimes can be arbitrarily more efficient than the LS and single quantile-based methods. The nice theoretical properties of CQR and WCQR have attracted increasing attention in recent years and applications of such a composite method can be found in various statistical problems, including regressions for correlated data (Tian et al., 2017, 2021), censored regression with measurement errors (Ma & Yin, 2011), time series regressions (Jiang et al., 2014), semi-parametric and nonparametric regressions (Kai

et al., 2011; Luo et al., 2019), and variable selection and feature screening procedures (Ma & Zhang, 2016; Xu, 2017). Though the parallel quantile curves are only guaranteed for models with error terms independent of the covariates, satisfactory performance of the WCQR has also been observed for heteroscedastic models of the form $y_i = m(\mathbf{x}_i) + \sigma(\mathbf{x}_i)\epsilon_i$, where $\sigma(\mathbf{x}_i)$ is the conditional scale (see Zhao et al. (2016); Jiang et al. (2014); Zhao et al. (2017) for heteroscedastic linear regression, Jiang et al. (2016a) and Jiang et al. (2016b) for heteroscedastic semiparametric models, and Kai et al. (2010); Guo et al. (2012); Sun et al. (2013); Huang & Zhan (2021) for heteroscedastic nonparametric models).

The existing literature has provided several insights to help understand CQR. Ma & Yin (2011) highlighted from a general modeling perspective that as CQR aims to find a set of parallel regression curves, it can be viewed as a compromise between a set of quantile regression curves with different slopes and intercepts and a single summary regression curve. Bradic et al. (2011) explained the efficiency of CQR from a nonparametric perspective – the quantile-specific check functions combined in CQR can be viewed as a set of basis functions to approximate the unknown log-likelihood function of the error distribution. Furthermore, CQR (argmins of the weighted average of quantile regression objective functions) also bears a close relationship with the L-estimator (weighted averages of argmins), a robust and efficient estimator that has a great estimation advantage for heterogeneous and asymmetric data (Koenker, 2005). Koenker (1984) proved that the optimal performance of CQR and L-estimator for linear regression is identical. For detailed discussions on the two methods, we refer readers to (Koenker, 2005) and (Bloznelis et al., 2019).

In this paper, we use the spatiotemporal mixed-effect model with random slopes and random intercepts to capture the heterogeneity in the costs data and assume the iid error terms to be independent of covariates in (1). Following the idea of CQR, the proposed ST-WCQR does not need extra assumptions since the independence between the error terms and covariates already ensures the parallel quantile regression curves. In the simulation studies, we show in Examples 1-3 that ST-WCQR exhibits better prediction performance and higher estimation efficiency than the conventional mean regression method and single quantile-based method for heterogeneous asymmetric errors (independent of the covariate) of the form $\epsilon_{ijk} = \sigma_{ij}\epsilon_{ijk}^*$, where $\sigma_{ij} \sim \text{Ga}(2, 2)$ allowing for spatiotemporal heterogeneity and ϵ_{ijk}^* are generated independently from one of the six symmetric or asymmetric distributions. In the following supplementary Section C.2, we will further verify the advantages of using ST-WCQR for heterogeneous errors that are correlated with covariates.

B Gibbs sampling algorithms

B.1 Gibbs sampling algorithm for the spatiotemporal weighted composite quantile regression

This section provides the Gibbs sampling algorithm for the proposed spatiotemporal weighted composite quantile regression model with shrinkage priors for the covariate effects. It is conducted by iteratively sampling from the following full conditional distributions of the unknown parameters given all the other parameters.

- **Update α_l :** For the τ_l -quantile ($l = 1, \dots, L$), suppose that the prior of α_l is assigned as $N(\mu_{0,\alpha_l}, \sigma_{0,\alpha_l})$. Then, the full conditional posterior distribution of α_l is a normal distribution with its mean and variance being

$$\mu_{\alpha_l} = \sigma_{\alpha_l} \left\{ \Sigma_{0,\alpha_l}^{-1} \mu_{0,\alpha_l} + \mathbf{1}_N^\top \mathbf{V}_l^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}_{\tilde{\Theta}} \tilde{\Theta} - \mathbf{Z}_\phi \phi - \mathbf{Z}_\psi \psi - \mathbf{Z}_{\tilde{\gamma}} \tilde{\gamma} - \xi_l \mathbf{v}_l) \right\}, \quad (\text{B.6})$$

and

$$\sigma_{\alpha_l} = \left(\Sigma_{0,\alpha_l}^{-1} + \mathbf{1}_N^\top \mathbf{V}_l^{-1} \mathbf{1}_N \right)^{-1}, \quad (\text{B.7})$$

where $\mathbf{1}_N$ is a $N \times 1$ vector of ones, \mathbf{V}_l is a $N \times N$ diagonal matrix of $\zeta_l \sigma_l v_{ijk,l}$, \mathbf{y} is a $N \times 1$ stacked response vector (first varying k and i then j), $\mathbf{y} = (y_{111}, \dots, y_{nJK_nJ})^\top$, $\mathbf{X} = (\mathbf{x}_{111}, \dots, \mathbf{x}_{nJK_nJ})^\top$ is a $p \times N$ design matrix, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown coefficients, $\tilde{\Theta}$ is the vectorized $n \times p$ random slope matrix $\Theta = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n)^\top = (\Theta_{\cdot 1}, \dots, \Theta_{\cdot p})$, and the corresponding design matrix for $\tilde{\Theta}$, the spatial effects ϕ , the temporal effects ψ and the spatio-temporal effects $\tilde{\gamma}$ are $\mathbf{Z}_{\tilde{\Theta}} = (\mathbf{Z}_{\Theta_1}, \dots, \mathbf{Z}_{\Theta_p})$, \mathbf{Z}_ϕ , \mathbf{Z}_ψ , and $\mathbf{Z}_{\tilde{\gamma}}$, respectively.

- **Update $\boldsymbol{\beta}$:** The full conditional posterior distribution of the vector of coefficients $\boldsymbol{\beta}$ is a p -dimensional multivariate normal distribution $N_p(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta)$, where

$$\boldsymbol{\mu}_\beta = \boldsymbol{\Sigma}_\beta \mathbf{X}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} (\mathbf{y} - \alpha_l \mathbf{1}_N - \mathbf{Z}_{\tilde{\Theta}} \tilde{\Theta} - \mathbf{Z}_\phi \phi - \mathbf{Z}_\psi \psi - \mathbf{Z}_{\tilde{\gamma}} \tilde{\gamma} - \xi_l \mathbf{v}_l), \quad (\text{B.8})$$

and

$$\boldsymbol{\Sigma}_\beta = \left(\boldsymbol{\Lambda}_\beta^{-2} / \tau_\beta^2 + \mathbf{X}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} \mathbf{X} \right)^{-1}. \quad (\text{B.9})$$

- **Update $\lambda_{\beta_h}^2$:** Update $\lambda_{\beta_h}^2$ from its $\text{IG}(1, 1/\eta_{\beta_h} + \beta_h^2/(2\tau_\beta^2))$ full conditional for all h .
- **Update η_{β_h} :** Update η_{β_h} from its $\text{IG}(1, 1 + 1/\lambda_{\beta_h}^2)$ full conditional for all h .

- **Update τ_{β}^2 :** Update τ_{β}^2 from its $\text{IG}((1+p)/2, 1/\eta_{\beta_0} + \beta^\top \Lambda_{\beta}^{-2} \beta/2)$ full conditional.
- **Update η_{β_0} :** The full conditional posterior distribution of $\lambda_{\beta_0}^2$ is $\text{IG}(1, 1 + 1/\tau_{\beta}^2)$.
- **Update $\Theta_{.h}$:** Assume the prior in (11). We update the h -th column of the random slope matrix, $\Theta_{.h}$, from its $N_n(\mu_{\Theta_h}, \Sigma_{\Theta_h})$ full conditional for $h = 1, \dots, p$, where

$$\mu_{\Theta_h} = \Sigma_{\Theta_h} \mathbf{Z}_{\Theta_h}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} (\mathbf{y} - \alpha_l \mathbf{1}_N - \mathbf{X}\beta - \mathbf{Z}_{\tilde{\Theta}_{.-h}} \tilde{\Theta}_{.-h} - \mathbf{Z}_{\phi} \phi - \mathbf{Z}_{\psi} \psi - \mathbf{Z}_{\tilde{\gamma}} \tilde{\gamma} - \xi_l \mathbf{v}_l), \quad (\text{B.10})$$

and

$$\Sigma_{\Theta_h} = \left(P / (\lambda_{\Theta_h}^2 \tau_{\Theta}^2) + \mathbf{Z}_{\Theta_h}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} \mathbf{Z}_{\Theta_h} \right)^{-1}. \quad (\text{B.11})$$

At the end of this step, the updated Θ_h is centered to ensure $\sum_{i=1}^n \Theta_{ih} = 0$ for all h . Though it is a mathematically informal way to impose the sum-zero constraints, it has recently been proved by Ferreira et al. (2021) that the conditional distribution obtained by using the improper ICAR prior and the ‘‘centering-on-the-fly’’ is equivalent to the sum-zero constrained ICAR distribution which is formally specified in Keefe et al. (2018). Given that the ‘‘centering-on-the-fly’’ is easy to implement numerically in the MCMC, we adopt this technique in our algorithm.

- **Update $\lambda_{\Theta_h}^2$:** Update $\lambda_{\Theta_h}^2$ from its $\text{IG}(n/2, 1/\eta_{\Theta_h} + \Theta_{.h}^\top \mathbf{P} \Theta_{.h} / (2\tau_{\Theta}^2))$ full conditional for all h .
- **Update η_{Θ_h} :** Update η_{Θ_h} from its $\text{IG}(1, 1 + 1/\lambda_{\Theta_h}^2)$ full conditional for all h .
- **Update τ_{Θ}^2 :** Update τ_{Θ}^2 from its $\text{IG}(n/2, 1/\eta_{\Theta_0} + \sum_{h=1}^p \Theta_{.h}^\top \mathbf{P} \Theta_{.h} / (2\lambda_{\Theta_h}^2))$ full conditional.
- **Update η_{Θ_0} :** Update η_{Θ_0} from its $\text{IG}(1, 1 + 1/\tau_{\Theta}^2)$ full conditional.
- **Update ϕ :** Given the prior in (5), we update ϕ from its $N_n(\mu_{\phi}, \Sigma_{\phi})$ full conditional, where

$$\mu_{\phi} = \Sigma_{\phi} \left\{ \mathbf{Z}_{\phi}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} (\mathbf{y} - \alpha_l \mathbf{1}_N - \mathbf{X}\beta - \mathbf{Z}_{\tilde{\Theta}} \tilde{\Theta} - \mathbf{Z}_{\psi} \psi - \mathbf{Z}_{\tilde{\gamma}} \tilde{\gamma} - \xi_l \mathbf{v}_l) \right\}, \quad (\text{B.12})$$

and

$$\Sigma_{\phi} = \left(\sigma_{\phi}^{-2} \mathbf{P} + \mathbf{Z}_{\phi}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} \mathbf{Z}_{\phi} \right)^{-1}. \quad (\text{B.13})$$

At the end of this step, the updated ϕ is centered to ensure $\sum_{i=1}^n \phi_i = 0$.

- **Update ψ :** Given the prior in (6), we update ψ from its $N_J(\boldsymbol{\mu}_\psi, \boldsymbol{\Sigma}_\psi)$ full conditional, where

$$\boldsymbol{\mu}_\psi = \boldsymbol{\Sigma}_\psi \left\{ \mathbf{Z}_\psi^\top \sum_{l=1}^L \mathbf{V}_l^{-1} (\mathbf{y} - \alpha_l \mathbf{1}_N - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}_{\tilde{\Theta}} \tilde{\Theta} - \mathbf{Z}_\phi \phi - \mathbf{Z}_{\tilde{\gamma}} \tilde{\gamma} - \xi_l \mathbf{v}_l) \right\}, \quad (\text{B.14})$$

and

$$\boldsymbol{\Sigma}_\psi = \left(\sigma_\phi^{-2} \mathbf{R} + \mathbf{Z}_\psi^\top \sum_{l=1}^L \mathbf{V}_l^{-1} \mathbf{Z}_\psi \right)^{-1}. \quad (\text{B.15})$$

At the end of this step, the updated ψ is centered to ensure $\sum_{j=1}^J \psi_j = 0$.

- **Update $\tilde{\gamma}$:** Given the prior in (7), the full conditional posterior distribution of $\tilde{\gamma}$ is given by $N_{nJ}(\boldsymbol{\mu}_{\tilde{\gamma}}, \boldsymbol{\Sigma}_{\tilde{\gamma}})$, where

$$\boldsymbol{\mu}_{\tilde{\gamma}} = \boldsymbol{\Sigma}_{\tilde{\gamma}} \left\{ \mathbf{Z}_{\tilde{\gamma}}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} (\mathbf{y} - \alpha_l \mathbf{1}_N - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}_{\tilde{\Theta}} \tilde{\Theta} - \mathbf{Z}_\phi \phi - \mathbf{Z}_\psi \psi - \xi_l \mathbf{v}_l) \right\}, \quad (\text{B.16})$$

and

$$\boldsymbol{\Sigma}_{\tilde{\gamma}} = \left(\sigma_{\tilde{\gamma}}^{-2} (\mathbf{R} \otimes \mathbf{P}) + \mathbf{Z}_{\tilde{\gamma}}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} \mathbf{Z}_{\tilde{\gamma}} \right)^{-1}. \quad (\text{B.17})$$

As the dimensions of \mathbf{R} and \mathbf{P} are low, we follow Neelon et al. (2015) to update the γ by year separately from the following full conditionals to speed up the Gibbs sampler. Recall that the j -th column of γ denoted by $\gamma_{\cdot j}$ is the spatiotemporal effects for the j -th year and denote $\mathbf{Z}_{\tilde{\gamma}} = (\mathbf{Z}_{\gamma_{\cdot 1}}, \dots, \mathbf{Z}_{\gamma_{\cdot m}})$. The full conditional posterior distribution of $\gamma_{\cdot j}$ is given by $N_n(\boldsymbol{\Sigma}_{\gamma_{\cdot j}} \boldsymbol{\mu}_{\gamma_{\cdot j}}, \boldsymbol{\Sigma}_{\gamma_{\cdot j}})$, where

$$\boldsymbol{\mu}_{\gamma_{\cdot j}} = \begin{cases} \sigma_{\tilde{\gamma}}^{-2} \mathbf{P} \gamma_{\cdot 2} + \mathbf{Z}_{\gamma_{\cdot 1}}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} (\mathbf{y} - \alpha_l \mathbf{1}_N - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}_{\tilde{\Theta}} \tilde{\Theta} - \mathbf{Z}_\phi \phi - \mathbf{Z}_\psi \psi - \mathbf{Z}_{\gamma_{\cdot 1}} \gamma_{\cdot 1} - \xi_l \mathbf{v}_l), & j = 1, \\ \sigma_{\tilde{\gamma}}^{-2} \mathbf{P} \gamma_{\cdot m-1} + \mathbf{Z}_{\gamma_{\cdot m}}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} (\mathbf{y} - \alpha_l \mathbf{1}_N - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}_{\tilde{\Theta}} \tilde{\Theta} - \mathbf{Z}_\phi \phi - \mathbf{Z}_\psi \psi - \mathbf{Z}_{\gamma_{\cdot m}} \gamma_{\cdot m} - \xi_l \mathbf{v}_l), & j = m, \\ \sigma_{\tilde{\gamma}}^{-2} \mathbf{P} (\gamma_{\cdot j-1} + \gamma_{\cdot j+1}) + \mathbf{Z}_{\gamma_{\cdot j}}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} (\mathbf{y} - \alpha_l \mathbf{1}_N - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}_{\tilde{\Theta}} \tilde{\Theta} - \mathbf{Z}_\phi \phi - \mathbf{Z}_\psi \psi - \mathbf{Z}_{\gamma_{\cdot j}} \gamma_{\cdot j} - \xi_l \mathbf{v}_l), & 1 < j < m, \end{cases} \quad (\text{B.18})$$

and

$$\boldsymbol{\Sigma}_{\gamma_{\cdot j}} = \begin{cases} \left(\sigma_{\tilde{\gamma}}^{-2} \mathbf{P} + \mathbf{Z}_{\gamma_{\cdot j}}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} \mathbf{Z}_{\gamma_{\cdot j}} \right)^{-1}, & j = 1, m, \\ \left(2\sigma_{\tilde{\gamma}}^{-2} \mathbf{P} + \mathbf{Z}_{\gamma_{\cdot j}}^\top \sum_{l=1}^L \mathbf{V}_l^{-1} \mathbf{Z}_{\gamma_{\cdot j}} \right)^{-1}, & j = 2, \dots, m-1, \end{cases} \quad (\text{B.19})$$

Then, update the $n \times J$ matrix γ according to these posterior samples. At the end of this step, the updated γ is centered by both row and column to ensure $\sum_{j=1}^J \gamma_{ij} = 0$ for all i and $\sum_{i=1}^n \gamma_{ij} = 0$ for all j .

- **Update σ_ϕ^2** : The derivation of the full conditional posterior distribution of $\sigma_\phi^2, \sigma_\psi^2$ and σ_γ^2 involves correctly specification of the exponents of these parameters in the corresponding ICAR priors in (5)-(7), for which we refer readers to the detailed discussion in Hodges et al. (2003); Rue & Held (2005); Keefe et al. (2018); Ferreira et al. (2021). Given the prior $\text{IG}(a, b)$, draw σ_ϕ^2 from its $\text{IG}(a^*, b^*)$ full conditional, where $a^* = a + (n - 1)/2$ and $b^* = b + \phi^\top \mathbf{P} \phi / 2$.
- **Update σ_ψ^2** : Given the prior $\text{IG}(c, d)$, draw σ_ψ^2 from its $\text{IG}(c^*, d^*)$ full conditional, where $c^* = c + (J - 1)/2$ and $d^* = d + \psi^\top \mathbf{R} \psi / 2$.
- **Update σ_γ^2** : Given the prior $\text{IG}(e, f)$, draw σ_γ^2 from its $\text{IG}(e^*, f^*)$ full conditional, where $e^* = e + (n - 1)(J - 1)/2$ and $f^* = f + \tilde{\gamma}^\top (\mathbf{R} \otimes \mathbf{P}) \tilde{\gamma} / 2$.
- **Update σ_l** : Given the prior $\text{IG}(g_l, h_l)$ for $l = 1, \dots, L$, draw σ_l from its $\text{IG}(g_l^*, h_l^*)$ full posterior distribution, where $g_l^* = g_l + 3N/2$ and $h_l^* = h_l + \sum_{i,j,k} \left\{ (y_{ijk} - \alpha_l - \mathbf{x}_{ijk}^\top (\boldsymbol{\beta} + \boldsymbol{\theta}_i) - \phi_i - \psi_j - \gamma_{ij})^2 / (2\zeta_l v_{ijk,l}) + v_{ijk,l} \right\}$.
- **Update $v_{ijk,l}$** : For all i, j, k, l , sample independently the inverse latent weights $v_{ijk,l}^{-1}$ from

$$v_{ijk,l}^{-1} \propto \text{InvGauss} \left(\frac{\sqrt{\xi_l^2 + 2\zeta_l}}{|y_{ijk} - \alpha_l - \mathbf{x}_{ijk}^\top (\boldsymbol{\beta} + \boldsymbol{\theta}_i) - \phi_i - \psi_j - \gamma_{ij}|}, \frac{\xi_l^2 + 2\zeta_l}{\zeta_l \sigma_l} \right). \quad (\text{B.20})$$

B.2 Gibbs sampling algorithm for the spatiotemporal mean regression

Recall that the linear mixed effects model (1) is given by

$$y_{ijk} = \mathbf{x}_{ijk}^\top (\boldsymbol{\beta} + \boldsymbol{\theta}_i) + \phi_i + \psi_j + \gamma_{ij} + \epsilon_{ijk}, \quad \forall i, j, k. \quad (\text{B.21})$$

where y_{ijk} and $\mathbf{x}_{ijk} \in \mathcal{R}^p$ are the response and the covariates of the k -th subject in region i and period j , $\boldsymbol{\beta}$ is the p -dimensional fixed effect, $\boldsymbol{\theta}_i$ is the random effect of the covariates for the i -th region, ϕ_i, ψ_j , and γ_{ij} are the unobservable random intercept for the region i and period j , and ϵ_{ijk} are the i.i.d. error terms of mean zero. We adopt the same notations as in the paper, i.e., $\mathbf{y} = (y_{111}, \dots, y_{nJK_{nJ}})^\top$, $\mathbf{X} = (\mathbf{x}_{111}, \dots, \mathbf{x}_{nJK_{nJ}})^\top$, $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n)^\top$, $\boldsymbol{\phi} = (\phi_1, \dots, \phi_n)^\top$, $\boldsymbol{\psi} = (\psi_1, \dots, \psi_J)^\top$, and the vectorized $n \times J$ matrix γ composed of

γ_{ij} is denoted as $\tilde{\gamma} = \text{vec}(\gamma) = (\gamma_{11}, \dots, \gamma_{nJ})^\top$. For the conventional mean regression method, the error terms are assumed to follow a normal distribution, i.e., $\epsilon_{ijk} \sim N(0, \sigma_\epsilon^2)$ (Lindley & Smith, 1972). Assign a $N(\mathbf{0}, 10^3 \mathbf{I})$ prior for the fixed effect β , an ICAR prior $\pi(\Theta_{.h} | \sigma_{\Theta_h}^2) \propto (\sigma_{\Theta_h}^2)^{-\frac{n-1}{2}} \exp\{-\Theta_{.h}^\top \mathbf{P} \Theta_{.h} / (2\sigma_{\Theta_h}^2)\}$ for the random effect θ_i for all i , the same noninformative $\text{IG}(0.001, 0.001)$ priors for σ_ϵ^2 and σ_{Θ_h} , and the same priors for the other parameters as specified in Section 4 and 5. Then, the corresponding Gibbs sampling algorithm for the proposed spatiotemporal mean regression model can be derived as follows.

- **Update β :** Given the prior $N_p(\mu_{0,\beta}, \Sigma_{0,\beta})$, the full conditional posterior distribution of the vector of fixed effects β is a p -dimensional multivariate normal distribution $N_p(\mu_\beta, \Sigma_\beta)$, where

$$\mu_\beta = \Sigma_\beta \left\{ \Sigma_{0,\beta}^{-1} \mu_{0,\beta} + \sigma_\epsilon^{-2} \mathbf{X}^\top (\mathbf{y} - \mathbf{Z}_{\tilde{\Theta}} \tilde{\Theta} - \mathbf{Z}_\phi \phi - \mathbf{Z}_\psi \psi - \mathbf{Z}_{\tilde{\gamma}} \tilde{\gamma}) \right\}, \quad (\text{B.22})$$

and

$$\Sigma_\beta = \left(\Sigma_{0,\beta}^{-1} + \sigma_\epsilon^{-2} \mathbf{X}^\top \mathbf{X} \right)^{-1}, \quad (\text{B.23})$$

where the corresponding design matrix for the vectorized random slope matrix Θ ($\tilde{\Theta}$), the spatial effects ϕ , the temporal effects ψ and the spatio-temporal effects $\tilde{\gamma}$ are $\mathbf{Z}_{\tilde{\Theta}} = (\mathbf{Z}_{\Theta_1}, \dots, \mathbf{Z}_{\Theta_p})$, \mathbf{Z}_ϕ , \mathbf{Z}_ψ , and $\mathbf{Z}_{\tilde{\gamma}}$, respectively.

- **Update $\Theta_{.h}$:** Assume the ICAR prior $\pi(\Theta_{.h} | \sigma_{\Theta_h}^2) \propto (\sigma_{\Theta_h}^2)^{-\frac{n-1}{2}} \exp\{-\Theta_{.h}^\top \mathbf{P} \Theta_{.h} / (2\sigma_{\Theta_h}^2)\}$. We update the h -th column of the random slope matrix, $\Theta_{.h}$, from its $N_n(\mu_{\Theta_h}, \Sigma_{\Theta_h})$ full conditional separately for $h = 1, \dots, p$, where

$$\mu_{\Theta_h} = \Sigma_{\Theta_h} \left\{ \sigma_\epsilon^{-2} \mathbf{Z}_{\Theta_h}^\top (\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}_{\tilde{\Theta}_{.-h}} \tilde{\Theta}_{.-h} - \mathbf{Z}_\phi \phi - \mathbf{Z}_\psi \psi - \mathbf{Z}_{\tilde{\gamma}} \tilde{\gamma}) \right\}, \quad (\text{B.24})$$

and

$$\Sigma_{\Theta_h} = \left(\sigma_{\Theta_h}^{-2} \mathbf{P} + \sigma_\epsilon^{-2} \mathbf{Z}_{\Theta_h}^\top \mathbf{Z}_{\Theta_h} \right)^{-1}. \quad (\text{B.25})$$

- **Update ϕ :** Given the prior in (7), update ϕ from its $N_n(\mu_\phi, \Sigma_\phi)$ full conditional, where

$$\mu_\phi = \Sigma_\phi \left\{ \sigma_\epsilon^{-2} \mathbf{Z}_\phi^\top (\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}_{\tilde{\Theta}} \tilde{\Theta} - \mathbf{Z}_\psi \psi - \mathbf{Z}_{\tilde{\gamma}} \tilde{\gamma}) \right\}, \quad (\text{B.26})$$

and

$$\Sigma_\phi = \left(\sigma_\phi^{-2} \mathbf{P} + \sigma_\epsilon^{-2} \mathbf{Z}_\phi^\top \mathbf{Z}_\phi \right)^{-1}. \quad (\text{B.27})$$

At the end of this step, the updated ϕ is centered to ensure $\sum_{i=1}^n \phi_i = 0$.

- **Update ψ :** Given the prior in (8), update ψ from its $N_J(\boldsymbol{\mu}_\psi, \boldsymbol{\Sigma}_\psi)$ full conditional, where

$$\boldsymbol{\mu}_\psi = \boldsymbol{\Sigma}_\psi \left\{ \sigma_\epsilon^{-2} \mathbf{Z}_\psi^\top (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}_{\tilde{\Theta}} \tilde{\boldsymbol{\Theta}} - \mathbf{Z}_\phi \boldsymbol{\phi} - \mathbf{Z}_{\tilde{\gamma}} \tilde{\boldsymbol{\gamma}}) \right\}, \quad (\text{B.28})$$

and

$$\boldsymbol{\Sigma}_\psi = \left(\sigma_\phi^{-2} \mathbf{R} + \sigma_\epsilon^{-2} \mathbf{Z}_\psi^\top \mathbf{Z}_\psi \right)^{-1}. \quad (\text{B.29})$$

At the end of this step, the updated ψ is centered to ensure $\sum_{j=1}^J \psi_j = 0$.

- **Update $\tilde{\boldsymbol{\gamma}}$:** Given the prior in (9), update $\tilde{\boldsymbol{\gamma}}$ from its $N_{nJ}(\boldsymbol{\mu}_{\tilde{\boldsymbol{\gamma}}}, \boldsymbol{\Sigma}_{\tilde{\boldsymbol{\gamma}}})$ full conditional, where

$$\boldsymbol{\mu}_{\tilde{\boldsymbol{\gamma}}} = \boldsymbol{\Sigma}_{\tilde{\boldsymbol{\gamma}}} \left\{ \sigma_\epsilon^{-2} \mathbf{Z}_{\tilde{\boldsymbol{\gamma}}}^\top (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}_{\tilde{\Theta}} \tilde{\boldsymbol{\Theta}} - \mathbf{Z}_\phi \boldsymbol{\phi} - \mathbf{Z}_\psi \boldsymbol{\psi}) \right\}, \quad (\text{B.30})$$

and

$$\boldsymbol{\Sigma}_{\tilde{\boldsymbol{\gamma}}} = \left(\sigma_{\tilde{\boldsymbol{\gamma}}}^{-2} (\mathbf{R} \otimes \mathbf{P}) + \sigma_\epsilon^{-2} \mathbf{Z}_{\tilde{\boldsymbol{\gamma}}}^\top \mathbf{Z}_{\tilde{\boldsymbol{\gamma}}} \right)^{-1}. \quad (\text{B.31})$$

Then, update the $n \times J$ matrix $\boldsymbol{\gamma}$ according to these posterior samples. In our algorithm, we update J columns of $\boldsymbol{\gamma}$ separately for computation efficiency. At the end of this step, the updated $\boldsymbol{\gamma}$ is centered by both row and column to ensure $\sum_{j=1}^J \gamma_{ij} = 0$ for all i and $\sum_{i=1}^n \gamma_{ij} = 0$ for all j .

- **Update $\sigma_{\Theta_h}^2$:** Given the prior $\text{IG}(v, w)$, draw $\sigma_{\Theta_h}^2$ from its $\text{IG}(v^*, w^*)$ full conditional, where $v^* = v + (n - 1)/2$ and $w^* = w + \boldsymbol{\Theta}_h^\top \mathbf{P} \boldsymbol{\Theta}_h / 2$.
- **Update σ_ϕ^2 :** Given the prior $\text{IG}(a, b)$, draw σ_ϕ^2 from its $\text{IG}(a^*, b^*)$ full conditional, where $a^* = a + (n - 1)/2$ and $b^* = b + \boldsymbol{\phi}^\top \mathbf{P} \boldsymbol{\phi} / 2$.
- **Update σ_ψ^2 :** Given the prior $\text{IG}(c, d)$, draw σ_ψ^2 from its $\text{IG}(c^*, d^*)$ full conditional, where $c^* = c + (J - 1)/2$ and $d^* = d + \boldsymbol{\psi}^\top \mathbf{R} \boldsymbol{\psi} / 2$.
- **Update $\sigma_{\tilde{\boldsymbol{\gamma}}}^2$:** Given the prior $\text{IG}(e, f)$, draw $\sigma_{\tilde{\boldsymbol{\gamma}}}^2$ from its $\text{IG}(e^*, f^*)$ full conditional, where $e^* = e + (n - 1)(J - 1)/2$ and $f^* = f + \tilde{\boldsymbol{\gamma}}^\top (\mathbf{R} \otimes \mathbf{P}) \tilde{\boldsymbol{\gamma}} / 2$.
- **Update σ_ϵ^2 :** Given the prior $\text{IG}(g, h)$, draw σ_ϵ^2 from its $\text{IG}(g^*, h^*)$ full posterior distribution, where $g^* = g + N/2$ and $h^* = h + \sum_{i,j,k} \{ y_{ijk} - \mathbf{x}_{ijk}^\top (\boldsymbol{\beta} + \boldsymbol{\theta}_i) - \phi_i - \psi_j - \gamma_{ij} \}^2 / 2$.

C Additional results for simulation studies

C.1 Additional results for Section 5

Tables C1-C6 provide additional simulation results for Section 5.

Table C1: The simulation results of ST-WCQR, STMR, STQR, and STQR_Neelon when the data sets are generated from symmetric error distributions for the very sparse case with $p = 20$ over 20 simulations. Optimal results are marked in bold.

ϵ	Example	Method	MSE(vary)	MSE(stat)	MSE		MAPE		β		γ		
					ϕ	ψ	θ	precision	recall	precision	recall	F1	F1
Norm	1	STQR_Neelon	-	<0.001	0.001	<0.001	0.002	0.505	0.952	1.000	0.973	-	-
		STQR	-	<0.001	0.001	<0.001	0.002	0.502	1.000	0.850	0.917	-	-
		STMR	-	<0.001	<0.001	0.001	0.507	0.905	1.000	0.945	-	-	-
	2	ST-WCQR	-	<0.001	<0.001	0.002	0.502	1.000	0.940	0.967	-	-	-
		ST-WCQR	-	<0.001	<0.001	0.003	0.503	1.000	0.990	0.994	-	-	-
		ST-WCQR	-	<0.001	<0.001	0.002	0.503	1.000	0.833	1.000	0.909	-	-
	3	STQR_Neelon	1.106	<0.001	0.003	<0.001	0.006	0.942	0.853	1.000	0.913	-	-
		STQR	<0.001	<0.001	0.001	<0.001	0.002	0.502	1.000	0.860	0.922	1.000	1.000
		STMR	0.001	<0.001	<0.001	0.002	0.510	0.913	1.000	0.950	0.783	1.000	0.860
t	1	ST-WCQR	-	<0.001	<0.001	0.002	0.502	1.000	0.960	0.978	1.000	1.000	
		ST-WCQR	-	<0.001	<0.001	0.002	0.502	1.000	0.990	0.994	1.000	1.000	
		ST-WCQR	-	<0.001	<0.001	0.002	0.503	1.000	1.000	1.000	1.000	1.000	
	2	STQR_Neelon	0.735	<0.001	<0.001	0.003	0.835	0.835	1.000	0.893	1.000	0.938	-
		STQR	<0.001	<0.001	0.001	0.384	1.000	0.800	0.899	1.000	0.899	1.000	1.000
		STMR	0.002	<0.001	0.001	0.397	1.000	0.625	1.000	0.769	0.726	1.000	0.814
	3	ST-WCQR	-	<0.001	0.001	0.002	0.483	1.000	0.947	0.970	1.000	1.000	1.000
		ST-WCQR	-	<0.001	<0.001	0.002	0.484	1.000	0.973	0.985	1.000	1.000	1.000
		ST-WCQR	-	<0.001	0.001	0.002	0.416	1.000	0.900	0.944	1.000	1.000	1.000
Cauchy	1	STQR_Neelon	-	<0.001	0.002	<0.001	0.003	0.585	0.947	1.000	0.968	-	-
		STQR	-	<0.001	0.002	0.003	0.584	1.000	0.830	0.906	-	-	
		STMR	-	<0.001	0.002	0.004	0.595	0.931	1.000	0.961	-	-	
	2	ST-WCQR	-	<0.001	0.002	<0.001	0.585	1.000	0.920	0.956	-	-	
		ST-WCQR	-	<0.001	0.001	<0.001	0.585	1.000	0.990	0.994	-	-	
		ST-WCQR	-	<0.001	0.002	0.002	0.585	1.000	1.000	1.000	1.000	1.000	
	3	STQR_Neelon	1.104	<0.001	0.004	0.002	0.008	1.048	0.862	1.000	0.920	-	-
		STQR	<0.001	<0.001	0.002	0.003	0.561	1.000	0.820	0.900	1.000	1.000	
		STMR	0.004	<0.001	0.002	0.004	0.603	0.900	1.000	0.944	0.792	1.000	0.870
Normal	1	ST-WCQR	-	<0.001	0.002	<0.001	0.003	0.562	1.000	0.980	0.989	1.000	1.000
		ST-WCQR	-	<0.001	0.002	<0.001	0.563	1.000	1.000	1.000	1.000	1.000	
		ST-WCQR	-	<0.001	0.002	0.003	0.563	1.000	1.000	1.000	1.000	1.000	
	2	STQR_Neelon	1.008	<0.001	0.004	0.002	0.006	1.026	0.857	1.000	0.919	-	-
		STQR	0.001	<0.001	0.002	0.003	0.583	1.000	0.850	0.917	1.000	1.000	
		STMR	0.004	<0.001	0.002	0.005	0.586	0.905	1.000	0.947	0.737	1.000	0.832
	3	ST-WCQR	-	<0.001	0.002	<0.001	0.584	1.000	0.930	0.961	1.000	1.000	
		ST-WCQR	-	<0.001	0.001	<0.001	0.585	1.000	0.990	0.994	1.000	1.000	
		ST-WCQR	-	<0.001	0.002	0.002	0.585	1.000	1.000	1.000	1.000	1.000	
Cauchy	1	STQR_Neelon	-	<0.001	0.002	<0.001	0.005	0.798	1.000	1.000	1.000	-	-
		STQR	-	<0.001	0.002	0.005	0.797	1.000	0.830	0.906	-	-	
		STMR	-	49.702	0.788	0.872	5.866	0.660	0.500	0.551	-	-	
	2	ST-WCQR	-	<0.001	0.003	<0.001	0.006	0.799	1.000	0.830	0.904	-	-
		ST-WCQR	-	<0.001	0.004	<0.001	0.006	0.799	1.000	0.810	0.890	-	-
		ST-WCQR	-	<0.001	0.005	0.001	0.799	1.000	0.870	0.926	-	-	
	3	STQR_Neelon	1.103	<0.001	0.005	0.002	0.011	1.336	1.000	1.000	1.000	-	-
		STQR	0.002	<0.001	0.002	0.006	0.773	1.000	0.830	0.906	1.000	1.000	
		STMR	29.160	18.200	0.977	3.078	6.529	0.762	0.480	0.550	0.050	0.050	
Normal	1	ST-WCQR	-	<0.001	0.003	0.001	0.006	0.774	1.000	0.800	0.884	1.000	1.000
		ST-WCQR	-	<0.001	0.003	0.001	0.007	0.774	1.000	0.800	0.879	1.000	1.000
		ST-WCQR	-	<0.001	0.003	0.002	0.008	0.774	1.000	0.800	0.879	1.000	1.000
	2	STQR_Neelon	1.005	<0.001	0.007	0.002	0.012	1.303	1.000	1.000	0.926	1.000	1.000
		STQR	0.002	<0.001	0.002	0.005	0.765	1.000	0.800	0.889	1.000	1.000	
		STMR	104.322	6.935	1.736	2.063	8.624	0.683	0.470	0.522	0.075	0.050	
	3	ST-WCQR	-	<0.001	0.002	<0.001	0.766	1.000	0.810	0.894	1.000	1.000	
		ST-WCQR	-	<0.001	0.003	<0.001	0.767	1.000	0.830	0.904	1.000	1.000	
		ST-WCQR	-	<0.001	0.003	<0.001	0.767	1.000	0.850	0.914	1.000	1.000	

Table C3: The simulation results of ST-WCQR, STMR, STQR, and STQR_Neelon when the data sets are generated from six error distributions for the dense case with $p = 20$ over 20 simulations. Optimal results are marked in bold.

ϵ	Example	Method	MSE(vary)	MSE(stat)	MSE		MAPE		β		γ		
					ϕ	ψ	θ	precision	recall	F1	precision	recall	F1
Norm	Ex1	STQR_Neelon	-	<0.001	0.001	<0.001	0.002	0.505	1.000	1.000	1.000	-	-
		STQR	-	<0.001	0.001	<0.001	0.501	1.000	1.000	1.000	1.000	-	-
		STMR	-	<0.001	<0.001	0.001	0.507	1.000	1.000	1.000	1.000	-	-
	Ex2	ST-WCQR	-	<0.001	<0.001	<0.001	0.502	1.000	1.000	1.000	1.000	-	-
		STQR	-	<0.001	<0.001	<0.001	0.502	1.000	1.000	1.000	1.000	-	-
		STMR	-	<0.001	<0.001	<0.001	0.502	1.000	1.000	1.000	1.000	-	-
	Ex3	STQR_Neelon	0.959	0.001	0.008	0.019	1.892	1.000	1.000	1.000	1.000	-	-
		STQR	<0.001	<0.001	0.002	0.003	0.485	1.000	1.000	1.000	1.000	1.000	1.000
		STMR	0.002	<0.001	<0.001	<0.001	0.493	1.000	1.000	1.000	1.000	0.955	1.000
t	Ex1	ST-WCQR	-	<0.001	0.001	<0.001	0.485	1.000	1.000	1.000	1.000	1.000	1.000
		STQR	-	<0.001	0.001	<0.001	0.486	1.000	1.000	1.000	1.000	1.000	1.000
		STMR	-	<0.001	0.001	<0.001	0.487	1.000	1.000	1.000	1.000	1.000	1.000
	Ex2	STQR_Neelon	1.032	0.002	0.013	0.003	0.020	1.943	1.000	1.000	1.000	-	-
		STQR	<0.001	<0.001	0.001	<0.001	0.002	0.504	1.000	1.000	1.000	1.000	1.000
		STMR	0.002	<0.001	0.001	<0.001	0.002	0.480	1.000	1.000	1.000	0.959	1.000
	Ex3	ST-WCQR	-	<0.001	0.001	<0.001	0.002	0.504	1.000	1.000	1.000	1.000	1.000
		STQR	-	<0.001	0.001	<0.001	0.002	0.505	1.000	1.000	1.000	1.000	1.000
		STMR	-	<0.001	0.001	<0.001	0.002	0.492	1.000	1.000	1.000	1.000	1.000
Cauchy	Ex1	STQR_Neelon	-	<0.001	0.002	<0.001	0.003	0.585	1.000	1.000	1.000	-	-
		STQR	-	<0.001	0.002	<0.001	0.003	0.583	1.000	1.000	1.000	-	-
		STMR	-	<0.001	0.002	0.001	0.004	0.595	1.000	1.000	1.000	-	-
	Ex2	ST-WCQR	-	<0.001	0.002	<0.001	0.003	0.584	1.000	1.000	1.000	-	-
		STQR	-	<0.001	0.001	<0.001	0.002	0.585	1.000	1.000	1.000	-	-
		STMR	-	<0.001	0.002	<0.001	0.003	0.568	1.000	1.000	1.000	1.000	1.000
	Ex3	ST-WCQR	-	<0.001	0.001	<0.001	0.003	0.569	1.000	1.000	1.000	1.000	0.980
		STQR	-	<0.001	0.001	<0.001	0.003	0.571	1.000	1.000	1.000	1.000	1.000
		STMR	-	<0.001	0.002	<0.001	0.003	0.567	1.000	1.000	1.000	1.000	1.000
Cauchy	Ex1	STQR_Neelon	1.032	0.002	0.015	0.003	2.046	1.000	1.000	1.000	-	-	
		STQR	0.001	<0.001	0.002	<0.001	0.005	0.561	1.000	1.000	1.000	1.000	1.000
		STMR	0.005	<0.001	0.002	0.001	0.004	0.573	1.000	1.000	1.000	0.966	0.979
	Ex2	ST-WCQR	-	<0.001	0.001	<0.001	0.004	0.562	1.000	1.000	1.000	1.000	1.000
		STQR	-	<0.001	0.002	<0.001	0.004	0.563	1.000	1.000	1.000	1.000	1.000
		STMR	-	<0.001	0.001	<0.001	0.004	0.567	1.000	1.000	1.000	1.000	1.000
	Ex3	STQR_Neelon	0.957	0.002	0.015	0.005	0.031	2.380	1.000	1.000	1.000	-	-
		STQR	0.002	<0.001	0.003	0.001	0.006	0.778	1.000	1.000	1.000	1.000	0.995
		STMR	146.144	23.868	0.791	0.943	0.877	6.226	0.850	0.195	0.286	0.271	0.045
Cauchy	Ex1	ST-WCQR	-	<0.001	0.003	0.006	0.798	1.000	1.000	1.000	1.000	0.995	0.997
		STQR	-	<0.001	0.004	<0.001	0.006	0.798	1.000	1.000	1.000	1.000	1.000
		STMR	-	<0.001	0.005	0.001	0.007	0.799	1.000	1.000	1.000	1.000	1.000
	Ex2	STQR_Neelon	0.957	0.002	0.015	0.005	0.031	2.380	1.000	1.000	1.000	-	-
		STQR	0.002	<0.001	0.003	0.001	0.006	0.778	1.000	1.000	1.000	1.000	0.995
		STMR	146.144	23.868	0.791	0.943	0.877	6.226	0.850	0.195	0.286	0.271	0.045
	Ex3	ST-WCQR	-	<0.001	0.004	0.007	0.783	1.000	1.000	1.000	1.000	1.000	0.995
		STQR	-	<0.001	0.004	0.007	0.783	1.000	1.000	1.000	1.000	1.000	0.995
		STMR	-	<0.001	0.004	0.007	0.783	1.000	1.000	1.000	1.000	1.000	0.997

Table C4: The simulation results of ST-WCQR, STMR, STQR, and STQR_Neelon when the data sets are generated from six error distributions for the very sparse case with $p = 8$ over 20 simulations. Optimal results are marked in bold.

ϵ	Example	Method	MSE(vary)	MSE(stat)	MSE		MAPE		β		γ		
					ϕ	ψ	θ	precision	recall	F1	precision	recall	F1
Norm	1	STQR_Neelon	-	<0.001	0.001	<0.001	0.002	0.505	0.967	1.000	0.980	-	-
		STQR	-	<0.001	0.001	<0.001	0.004	0.505	1.000	1.000	1.000	-	-
		STMR	-	<0.001	<0.001	0.001	0.001	0.506	0.983	1.000	0.990	-	-
	2	ST-WCQR	-	<0.001	0.001	<0.001	0.003	0.505	0.983	1.000	0.990	-	-
		ST-WCQR	-	<0.001	0.001	<0.001	0.003	0.505	0.967	1.000	0.980	-	-
		STQR_Neelon	1.003	<0.001	0.002	<0.001	0.003	0.506	0.967	1.000	0.980	-	-
	3	STQR	<0.001	<0.001	0.001	<0.001	0.003	0.503	1.000	1.000	1.000	1.000	1.000
		STMR	0.001	<0.001	<0.001	0.002	0.505	0.983	1.000	0.990	0.804	1.000	0.862
		ST-WCQR	<0.001	<0.001	0.001	<0.001	0.002	0.504	0.983	1.000	0.990	1.000	1.000
t	1	STQR_Neelon	0.666	<0.001	0.001	<0.001	0.002	0.744	0.892	1.000	0.933	-	-
		STQR	<0.001	<0.001	0.001	<0.001	0.003	0.509	1.000	1.000	1.000	1.000	1.000
		STMR	<0.001	<0.001	<0.001	0.002	0.509	0.527	0.942	1.000	0.963	0.867	1.000
	2	ST-WCQR	<0.001	<0.001	<0.001	<0.001	0.003	0.509	0.967	1.000	0.980	1.000	1.000
		ST-WCQR	<0.001	<0.001	<0.001	<0.001	0.003	0.510	0.950	1.000	0.970	1.000	1.000
		STQR_Neelon	<0.001	<0.001	<0.001	0.002	0.510	0.510	0.917	1.000	0.950	1.000	1.000
	3	STQR	-	<0.001	0.002	<0.001	0.003	0.586	0.983	1.000	0.990	-	-
		STMR	-	<0.001	0.002	<0.001	0.004	0.586	1.000	1.000	1.000	-	-
		STMR	-	<0.001	0.002	0.001	0.004	0.591	0.892	1.000	0.930	-	-
Cauchy	1	STQR_Neelon	1.003	<0.001	0.002	<0.001	0.006	0.852	0.917	1.000	0.950	-	-
		STQR	<0.001	<0.001	0.002	<0.001	0.003	0.586	1.000	1.000	1.000	1.000	1.000
		STMR	0.004	<0.001	0.002	<0.001	0.005	0.591	0.933	1.000	0.960	0.867	1.000
	2	ST-WCQR	0.001	<0.001	0.002	<0.001	0.002	0.586	1.000	1.000	1.000	1.000	1.000
		ST-WCQR	0.001	<0.001	0.001	<0.001	0.002	0.586	0.983	1.000	0.990	1.000	1.000
		STQR_Neelon	0.664	<0.001	0.001	<0.001	0.002	0.839	1.000	1.000	1.000	1.000	1.000
	3	STQR	<0.001	<0.001	0.001	<0.001	0.003	0.593	1.000	1.000	1.000	1.000	1.000
		STMR	0.003	<0.001	0.003	0.001	0.005	0.615	0.920	1.000	0.949	0.800	1.000
		STMR	<0.001	<0.001	0.001	<0.001	0.003	0.593	0.983	1.000	0.990	1.000	1.000
Gamma	1	STQR_Neelon	-	<0.001	0.001	<0.001	0.002	0.800	1.000	1.000	1.000	-	-
		STQR	-	<0.001	0.002	<0.001	0.005	0.799	1.000	1.000	1.000	-	-
		STMR	-	51.909	0.838	1.006	0.816	4.213	0.475	0.300	0.358	-	-
	2	ST-WCQR	-	<0.001	0.003	<0.001	0.006	0.800	1.000	1.000	1.000	-	-
		ST-WCQR	-	<0.001	0.004	<0.001	0.006	0.800	0.983	1.000	0.990	-	-
		STQR_Neelon	1.003	<0.001	0.005	0.001	0.007	0.801	0.983	1.000	0.990	-	-
	3	STQR	0.001	<0.001	0.004	0.001	0.008	1.122	1.000	1.000	1.000	-	-
		STMR	7.095	<0.001	0.002	<0.001	0.005	0.798	1.000	1.000	1.000	1.000	1.000
		STMR	0.002	3.721	1.218	1.026	1.041	3.730	0.433	0.300	0.340	0.050	0.050
Normal	1	STQR_Neelon	0.664	<0.001	0.003	0.001	0.007	1.095	1.000	1.000	1.000	-	-
		STQR	0.001	<0.001	0.002	<0.001	0.005	0.832	1.000	1.000	1.000	1.000	1.000
		STMR	18.795	17.083	1.959	1.772	1.465	4.230	0.450	0.350	0.383	0.050	0.050
	2	ST-WCQR	0.002	<0.001	0.002	0.001	0.006	0.834	1.000	1.000	1.000	1.000	1.000
		ST-WCQR	0.002	<0.001	0.003	0.001	0.008	0.834	1.000	1.000	1.000	1.000	1.000
		STQR_Neelon	0.664	<0.001	0.003	0.001	0.009	0.835	0.983	1.000	0.990	1.000	1.000
	3	STQR	0.001	<0.001	0.002	<0.001	0.005	0.832	1.000	1.000	1.000	1.000	1.000
		STMR	18.795	17.083	1.959	1.772	1.465	4.230	0.450	0.350	0.383	0.050	0.050
		ST-WCQR	0.002	<0.001	0.003	0.001	0.008	0.834	1.000	1.000	1.000	1.000	1.000

Table C5: The simulation results of ST-WCQR, STMR, STQR, and STQR_Neelon when the data sets are generated from six error distributions for the sparse case with $p = 8$ over 20 simulations. Optimal results are marked in bold.

ϵ	Example	Method	MSE(vary)	MSE(stat)	MSE		MAPE		β		γ			
					ψ	ϕ	θ	precision	recall	precision	recall	F1	F1	
Norm	1	STQR_Neelon	-	<0.001	0.001	<0.001	0.002	0.505	0.980	1.000	0.989	-	-	
		STQR	-	<0.001	0.001	<0.001	0.002	0.504	1.000	0.925	0.957	-	-	
		STMR	-	<0.001	<0.001	<0.001	0.001	0.506	0.990	1.000	0.994	-	-	
	2	ST-WCQR	$L = 3$	-	<0.001	<0.001	0.002	0.504	1.000	0.888	0.936	-	-	
		ST-WCQR	$L = 5$	-	<0.001	<0.001	0.002	0.504	1.000	0.900	0.943	-	-	
		ST-WCQR	$L = 9$	-	<0.001	<0.001	0.002	0.504	1.000	0.863	0.921	-	-	
	3	STQR_Neelon	0.919	<0.001	0.003	<0.001	0.006	0.893	0.963	1.000	0.979	-	-	
		STQR	<0.001	<0.001	0.001	<0.001	0.002	0.486	1.000	1.000	1.000	1.000	1.000	
		STMR	0.001	<0.001	<0.001	<0.001	0.001	0.507	0.990	1.000	0.994	0.867	1.000	
t	1	ST-WCQR	$L = 3$	<0.001	0.001	<0.001	0.002	0.485	1.000	0.888	0.933	1.000	1.000	
		ST-WCQR	$L = 5$	<0.001	0.001	<0.001	0.002	0.485	1.000	0.913	0.950	1.000	1.000	
		ST-WCQR	$L = 9$	<0.001	0.001	<0.001	0.002	0.486	1.000	0.913	0.950	1.000	1.000	
	2	STQR_Neelon	0.916	<0.001	0.002	<0.001	0.003	0.586	0.990	1.000	0.994	-	-	
		STQR	0.001	<0.001	0.001	<0.001	0.003	0.484	1.000	0.875	0.929	1.000	1.000	
		STMR	0.004	<0.001	<0.001	<0.001	0.001	0.507	0.990	1.000	0.994	0.867	1.000	
	3	ST-WCQR	$L = 3$	-	<0.001	0.001	<0.001	0.002	0.485	1.000	0.925	0.957	-	-
		ST-WCQR	$L = 5$	-	<0.001	0.001	<0.001	0.002	0.485	1.000	0.950	0.971	-	-
		ST-WCQR	$L = 9$	-	<0.001	0.001	<0.001	0.002	0.486	1.000	0.938	0.964	-	-
Cauchy	1	STQR_Neelon	0.915	<0.001	0.002	<0.001	0.007	0.994	0.970	1.000	0.983	-	-	
		STQR	0.001	<0.001	0.002	<0.001	0.003	0.563	1.000	1.000	1.000	1.000	1.000	
		STMR	0.004	<0.001	0.002	0.001	0.004	0.594	0.970	1.000	0.983	0.950	1.000	
	2	ST-WCQR	$L = 3$	-	<0.001	0.002	<0.001	0.002	0.587	1.000	0.925	0.957	1.000	1.000
		ST-WCQR	$L = 5$	-	<0.001	0.001	<0.001	0.002	0.587	1.000	0.950	0.971	1.000	1.000
		ST-WCQR	$L = 9$	-	<0.001	0.001	<0.001	0.002	0.587	1.000	0.925	0.957	1.000	1.000
	3	STQR_Neelon	0.915	<0.001	0.004	0.002	0.011	1.265	1.000	1.000	1.000	-	-	
		STQR	0.002	<0.001	0.002	<0.001	0.006	0.775	1.000	1.000	1.000	1.000	1.000	
		STMR	31.235	20.599	0.797	1.003	0.840	4.196	0.846	0.550	0.642	0.200	0.133	
Normal	1	ST-WCQR	$L = 3$	<0.001	0.003	<0.001	0.006	0.800	1.000	0.963	0.979	-	-	
		ST-WCQR	$L = 5$	<0.001	0.003	<0.001	0.006	0.799	1.000	0.988	0.993	-	-	
		ST-WCQR	$L = 9$	<0.001	0.003	0.001	0.007	0.801	1.000	0.925	0.957	-	-	
	2	STQR_Neelon	0.915	<0.001	0.004	0.002	0.011	1.265	1.000	1.000	1.000	-	-	
		STQR	0.002	<0.001	0.002	<0.001	0.006	0.775	1.000	1.000	1.000	1.000	1.000	
		STMR	31.235	20.599	0.797	1.003	0.840	4.196	0.846	0.550	0.642	0.200	0.133	
	3	ST-WCQR	$L = 3$	0.003	<0.001	0.003	0.006	0.777	1.000	1.000	1.000	1.000	1.000	
		ST-WCQR	$L = 5$	0.003	<0.001	0.003	0.006	0.777	1.000	1.000	1.000	1.000	1.000	
		ST-WCQR	$L = 9$	0.003	<0.001	0.003	0.002	0.777	1.000	1.000	1.000	1.000	1.000	

Table C6: The simulation results of ST-WCQR, STMR, STQR, and STQR_Neelon when the data sets are generated from six error distributions for the dense case with $p = 8$ over 20 simulations. Optimal results are marked in bold.

ϵ	Example	Method	MSE ^(vary)	MSE ^(stat)	MSE		MAPE	β		γ	
					ϕ	ψ		precision	recall	precision	recall
Norm	1	STQR_Neelon	-	<0.001	0.001	<0.001	0.002	1.000	1.000	1.000	-
		STQR	-	<0.001	0.001	<0.001	0.002	1.000	1.000	1.000	-
		STMR	-	<0.001	<0.001	0.001	0.506	1.000	1.000	1.000	-
	2	ST-WCQR	-	<0.001	<0.001	<0.001	0.504	1.000	1.000	1.000	-
		STQR_Neelon	-	<0.001	<0.001	<0.001	0.504	1.000	1.000	1.000	-
		STQR	0.945	<0.001	0.005	0.008	1.242	1.000	1.000	1.000	-
	3	STQR	<0.001	<0.001	0.002	0.002	0.487	1.000	1.000	1.000	1.000
		STMR	0.001	<0.001	<0.001	0.001	0.507	1.000	1.000	0.933	1.000
		ST-WCQR	0.001	<0.001	<0.001	<0.001	0.488	1.000	1.000	1.000	1.000
t	1	STQR_Neelon	1.004	<0.001	0.004	0.002	0.009	1.000	1.000	1.000	-
		STQR	<0.001	<0.001	0.001	<0.001	0.003	0.476	1.000	1.000	1.000
		STMR	0.001	<0.001	<0.001	<0.001	0.002	0.480	1.000	1.000	0.933
	2	ST-WCQR	<0.001	<0.001	0.001	<0.001	0.002	0.477	1.000	1.000	1.000
		STQR_Neelon	<0.001	<0.001	<0.001	<0.001	0.002	0.477	1.000	1.000	1.000
		STQR	<0.001	<0.001	<0.001	<0.001	0.002	0.478	1.000	1.000	1.000
	3	STQR	-	<0.001	0.002	<0.001	0.003	0.586	1.000	1.000	-
		STMR	-	<0.001	0.002	<0.001	0.003	0.586	1.000	1.000	-
		ST-WCQR	-	<0.001	0.002	<0.001	0.004	0.592	1.000	1.000	-
Cauchy	1	STQR_Neelon	0.943	<0.001	0.006	0.002	0.009	1.000	1.000	1.000	-
		STQR	0.001	<0.001	0.002	<0.001	0.003	0.585	1.000	1.000	1.000
		STMR	0.004	<0.001	0.002	0.001	0.004	0.595	1.000	1.000	0.980
	2	ST-WCQR	0.001	<0.001	0.002	<0.001	0.002	0.587	1.000	1.000	1.000
		STQR_Neelon	0.002	<0.001	0.001	<0.001	0.002	0.586	1.000	1.000	1.000
		STQR	1.003	<0.001	0.001	<0.001	0.002	0.587	1.000	1.000	1.000
	3	STQR	0.001	<0.001	0.002	<0.001	0.003	0.558	1.000	1.000	1.000
		STMR	0.004	<0.001	0.002	0.001	0.005	0.566	1.000	1.000	0.990
		ST-WCQR	0.001	<0.001	0.002	<0.001	0.002	0.558	1.000	1.000	1.000
1	1	STQR_Neelon	-	<0.001	0.001	<0.001	0.002	1.000	1.000	1.000	-
		STQR	-	<0.001	0.002	<0.001	0.005	0.798	1.000	1.000	-
		STMR	-	14.117	0.796	1.001	0.841	4.089	0.750	0.194	0.281
	2	ST-WCQR	-	<0.001	0.004	<0.001	0.006	0.800	1.000	1.000	-
		STQR_Neelon	-	<0.001	0.004	<0.001	0.006	0.800	1.000	1.000	-
		STQR	0.942	0.001	0.005	0.001	0.007	0.801	1.000	1.000	-
	3	STQR	0.002	<0.001	0.009	0.002	0.016	1.666	1.000	1.000	-
		STMR	26.508	<0.001	0.003	<0.001	0.005	0.799	1.000	1.000	1.000
		ST-WCQR	0.003	<0.001	0.004	<0.001	0.006	0.801	1.000	1.000	0.225
2	1	STQR_Neelon	1.001	<0.001	0.005	0.001	0.006	0.801	1.000	1.000	1.000
		STQR	0.002	<0.001	0.002	<0.001	0.003	0.758	1.000	1.000	1.000
		STMR	6.937	11.905	1.052	0.832	1.046	4.031	0.700	0.200	0.289
	2	ST-WCQR	0.003	<0.001	0.003	<0.001	0.005	0.759	1.000	1.000	1.000
		STQR_Neelon	0.003	<0.001	0.004	<0.001	0.005	0.760	1.000	1.000	1.000
		STQR	1.001	0.001	0.009	0.003	0.018	1.695	1.000	1.000	1.000
	3	STQR	0.002	<0.001	0.002	<0.001	0.005	0.758	1.000	1.000	1.000
		STMR	0.003	<0.001	0.003	<0.001	0.005	0.759	1.000	1.000	0.013
		ST-WCQR	0.003	<0.001	0.004	<0.001	0.006	0.761	1.000	1.000	1.000

C.2 Additional simulation studies

Example 4. (*Heterogeneous random errors*) In this example, we consider the case when the error terms are correlated with covariates for $i = 1, \dots, 7; j = 1, 2, 3; k = 1, \dots, 500$:

$$y_{ijk} = \mathbf{x}_{ijk}(\boldsymbol{\beta} + \boldsymbol{\theta}_i) + \phi_i + \psi_j + \gamma_{ij} + (1 + x_{ijk,1})\epsilon_{ijk}, \quad (\text{C.32})$$

where $\boldsymbol{\beta} = (1, -2, \dots, (-1)^{p/4-1}p/4, 0, \dots, 0)^\top$ and $p = 20$. Other settings are the same as those in Examples 1 and 2.

Table C7 compares the results of ST-WCQR, STMR, STQR, and STQR_Neelon (Neelon et al., 2015) for Example 4 over 20 simulations. We find that ST-WCQR has better prediction performance than other methods in most situations for heteroscedastic models. Even for cases when a better method exists, ST-WCQR performs nearly as well as the superior method. Similar to the case in Examples 2 & 3, STQR_Neelon fails to select and estimate the spatially varying effects as other methods do and thus has the largest prediction errors for the heterogeneous data. STMR performs consistently the worst especially when the error distribution is asymmetric or has infinite variance like Cauchy distribution.

Table C7: The simulation results of ST-WCQR, STMR, STQR, and STQR_Neelon in Example 4 when the data sets are generated from six error distributions for the very sparse case with $p = 20$ over 20 simulations. Optimal results are marked in bold.

ϵ	Method	MSE			MAPE			β			γ		
		MSE(vary)	MSE(stat)	ψ	ϕ	θ	precision	recall	F1	precision	recall	F1	
Norm	STQR_Neelon	0.917	<0.001	0.002	0.041	0.013	0.829	1.000	0.934	-	-	-	-
	STQR	<0.001	<0.001	<0.001	<0.001	<0.001	0.406	1.000	0.889	1.000	1.000	1.000	1.000
	STMR	0.003	<0.001	<0.001	0.002	0.004	0.416	0.855	1.000	0.915	0.757	1.000	0.844
ST-WCQR	$L = 3$	<0.001	<0.001	<0.001	<0.001	0.001	0.405	1.000	0.950	1.000	1.000	1.000	1.000
	$L = 5$	<0.001	<0.001	<0.001	<0.001	0.001	0.406	1.000	0.978	1.000	1.000	1.000	1.000
	$L = 9$	0.001	<0.001	<0.001	<0.001	0.001	0.405	1.000	0.994	1.000	1.000	1.000	1.000
t	STQR_Neelon	0.918	<0.001	0.002	0.050	0.012	0.916	1.000	0.949	-	-	-	-
	STQR	<0.001	<0.001	<0.001	<0.001	<0.001	0.472	1.000	0.894	1.000	1.000	1.000	1.000
	STMR	0.009	0.001	0.002	0.005	0.011	0.511	0.894	1.000	0.940	0.867	1.000	0.920
Cauchy	$L = 3$	0.001	<0.001	<0.001	<0.001	0.001	0.472	1.000	0.950	1.000	1.000	1.000	1.000
	$L = 5$	0.001	<0.001	<0.001	0.001	0.001	0.472	1.000	0.961	1.000	1.000	1.000	1.000
	$L = 9$	0.001	<0.001	<0.001	0.001	0.002	0.473	1.000	1.000	1.000	1.000	1.000	1.000
LN	STQR_Neelon	5.359	<0.001	0.002	0.056	0.012	1.157	1.000	1.000	-	-	-	-
	STQR	4.445	<0.001	<0.001	0.002	0.002	0.654	1.000	0.949	1.000	1.000	1.000	1.000
	STMR	18.982	12.570	0.982	1.237	2.286	5.172	0.646	0.380	0.465	<0.001	<0.001	<0.001
χ^2	$L = 3$	4.510	<0.001	<0.001	0.002	0.003	0.657	1.000	0.950	1.000	1.000	1.000	1.000
	$L = 5$	4.527	<0.001	<0.001	0.003	0.004	0.657	1.000	0.880	0.931	1.000	1.000	1.000
	$L = 9$	4.537	<0.001	<0.001	0.003	0.004	0.659	1.000	0.880	0.931	1.000	1.000	1.000
Ga	STQR_Neelon	1.149	<0.001	0.026	0.122	0.168	0.910	0.975	1.000	0.986	-	-	-
	STQR	0.379	<0.001	0.014	0.116	0.089	0.487	0.970	0.970	1.000	1.000	1.000	1.000
	STMR	0.199	0.003	0.109	0.371	0.694	0.928	0.884	1.000	0.933	0.875	1.000	0.923
ST-WCQR	$L = 3$	0.370	<0.001	0.012	0.082	0.089	0.485	0.980	0.980	1.000	1.000	1.000	1.000
	$L = 5$	0.375	<0.001	0.012	0.071	0.089	0.487	0.990	0.990	1.000	1.000	1.000	1.000
	$L = 9$	0.375	<0.001	0.012	0.064	0.087	0.487	0.990	0.990	1.000	1.000	1.000	1.000
χ^2	STQR_Neelon	1.228	<0.001	0.038	0.175	0.225	1.151	0.934	0.980	0.951	-	-	-
	STQR	0.484	<0.001	0.021	0.216	0.220	0.720	0.990	0.990	0.990	1.000	1.000	1.000
	STMR	0.265	0.002	0.169	0.513	1.033	1.134	0.921	1.000	0.956	0.892	1.000	0.933
ST-WCQR	$L = 3$	0.461	<0.001	0.020	0.156	0.145	0.723	0.990	0.990	1.000	1.000	1.000	1.000
	$L = 5$	0.466	<0.001	0.020	0.135	0.143	0.719	1.000	1.000	1.000	1.000	1.000	1.000
	$L = 9$	0.467	<0.001	0.020	0.119	0.141	0.720	1.000	1.000	1.000	1.000	1.000	1.000
ST-WCQR	STQR_Neelon	0.948	<0.001	0.021	0.088	0.135	0.728	0.942	1.000	0.968	-	-	-
	STQR	0.085	<0.001	0.010	0.079	0.073	0.328	1.000	1.000	1.000	1.000	1.000	1.000
	STMR	0.067	<0.001	0.041	0.131	0.257	0.462	0.871	1.000	0.929	0.768	1.000	0.847
ST-WCQR	$L = 3$	0.076	<0.001	0.011	0.067	0.077	0.334	0.975	1.000	0.986	1.000	1.000	1.000
	$L = 5$	0.074	<0.001	0.011	0.062	0.078	0.338	0.958	1.000	0.977	1.000	1.000	1.000
	$L = 9$	0.073	<0.001	0.011	0.058	0.079	0.339	0.908	1.000	0.950	1.000	1.000	1.000

D Additional results for the case study

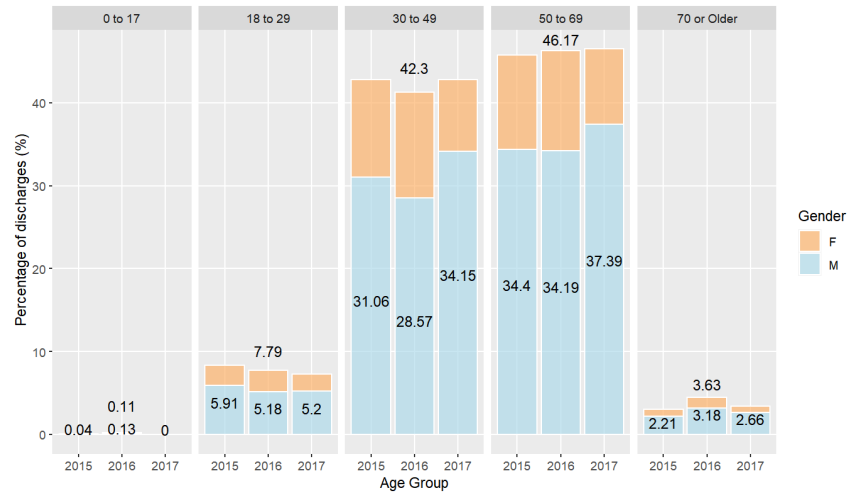


Figure D.1: The yearly age-gender composition of the discharged patients admitted with alcohol-related disorders from 2015 to 2017. For each year, the percentage of male inpatients' discharge records of each age group among all the records of that year is labeled in the middle of each bar. The numbers displayed at the top of the bars are average percentages of discharge for each age group over the three years.

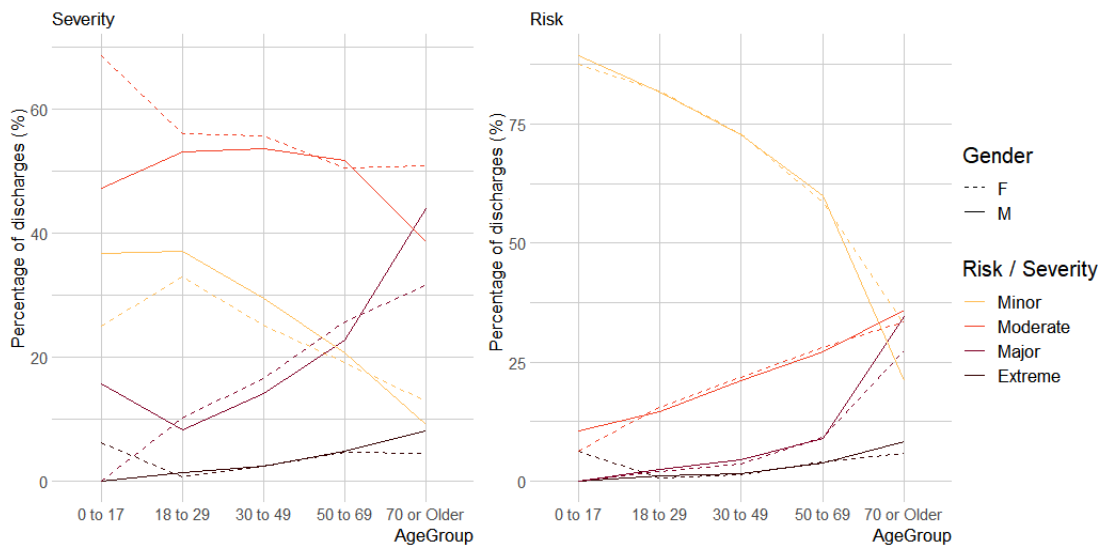


Figure D.2: The composition of the discharged patients admitted with alcohol-related disorders by the severity of illness (left panel) and the risk of mortality (right panel) given age and gender. The severity of illness (the degree of physiologic decompensation or organ system derangement) and the risk of mortality (the likelihood of dying) are four-level measures assessed through a uniform set of diagnosis-based methods in the All Patient Refined Diagnosis Related Groups payment system used by many US hospitals for inpatient visit classification (Averill et al., 2003).

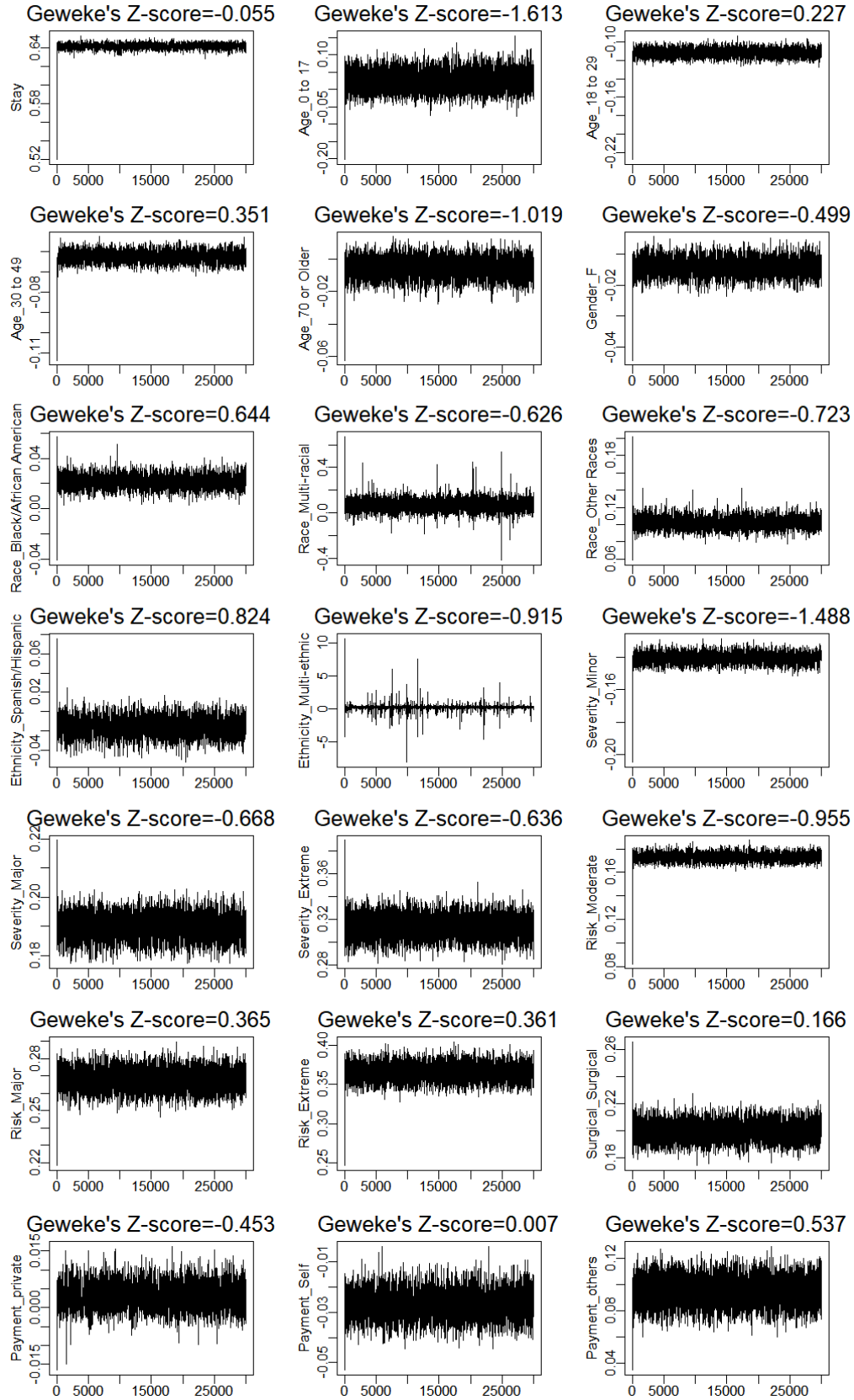


Figure D.3: The trace plots of the fixed effects estimated by the ST-WCQR with $L_{opt} = 9$ and their Geweke's z-scores.

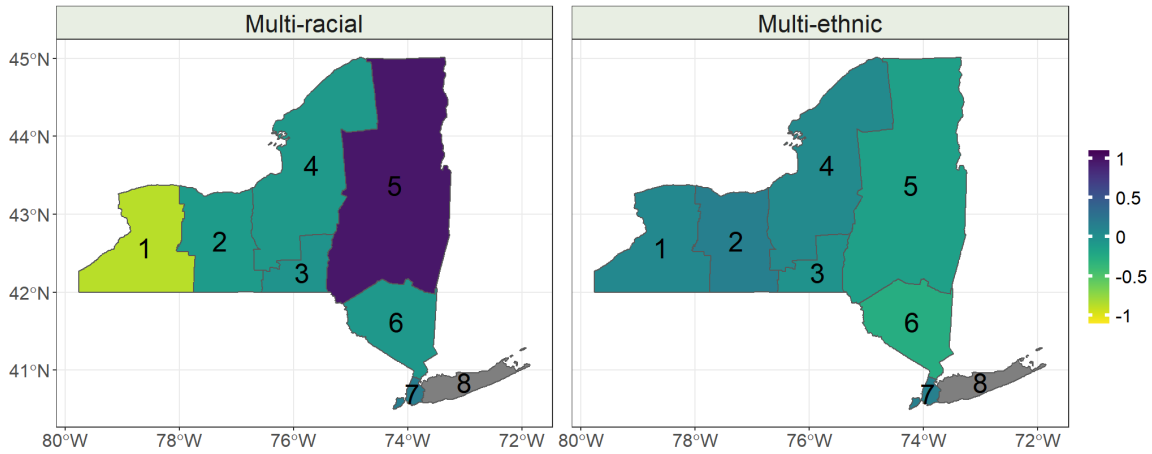


Figure D.4: The plot of the estimated significant spatially varying coefficients θ_i for the seven healthcare service areas by ST-WCQR. We label the areas as 1: Western NY, 2: Finger Lakes, 3: Southern Tier, 4: CNY, 5: Capital/Adirondack, 6: Hudson Valley, 7: NYC.

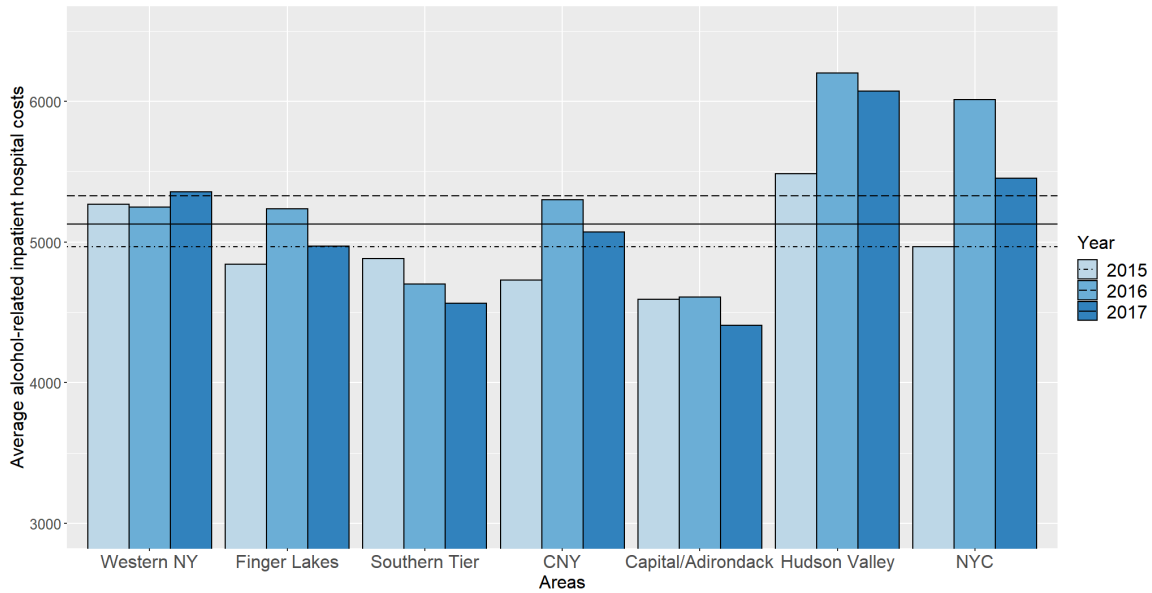


Figure D.5: The estimated area-specific average inpatient hospital costs for the reference group (grouped barplot) from 2015 to 2017 and the estimated yearly statewide average costs (horizontal lines). The reference group is non-Hispanic, white, male patients between the age of 50-69 with federal insurance, who receive medical treatment with an average length of stay (4.9 days) for minor risk of mortality and moderate severity of illness, and who have no spatiotemporal effects.

References

- Averill, R. F., Goldfield, N., Hughes, J. S., Bonazelli, J., McCullough, E. C., Steinbeck, B. A., Mullin, R., Tang, A. M., Muldoon, J., Turner, L., & Gay, J. (2003). All patient refined diagnosis related groups (APR-DRGs): Methodology overview, . Available at <https://www.hcup-s.ahrq.gov/db/nation/nis/APR-DRGsV20MethodologyOverviewandBibliography.pdf> (Accessed March 21, 2022).
- Bloznelis, D., Claeskens, G., & Zhou, J. (2019). Composite versus model-averaged quantile regression. *J. Stat. Plan. Inference*, *200*, 32–46.
- Bradic, J., Fan, J., & Wang, W. (2011). Penalized composite quasi-likelihood for ultrahigh dimensional variable selection. *J. R. Stat. Soc., B: Stat. Methodol.*, *73*, 325–349.
- Ferreira, M. A. R., Porter, E. M., & Franck, C. T. (2021). Fast and scalable computations for Gaussian hierarchical models with intrinsic conditional autoregressive spatial random effects. *Comput. Stat. Data Anal.*, *162*, 107264.
- Guo, J., Tian, M., & Zhu, K. (2012). New efficient and robust estimation in varying-coefficient models with heteroscedasticity. *Stat. Sin.*, *22*, 1075–1101.
- Hodges, J. S., Carlin, B. P., & Fan, Q. (2003). On the Precision of the Conditionally Autoregressive Prior in Spatial Models. *Biometrics*, *59*, 317–322.
- Huang, X., & Zhan, Z. (2021). Local composite quantile regression for regression discontinuity. *J. Bus. Econ. Stat.*, *0*, 1–13.
- Jiang, J., Jiang, X., & Song, X. (2014). Weighted composite quantile regression estimation of dtarch models. *Econom. J.*, *17*, 1–23.
- Jiang, R., Qian, W.-M., & Zhou, Z.-G. (2016a). Single-index composite quantile regression with heteroscedasticity and general error distributions. *Stat. Pap.*, *57*, 185–203.
- Jiang, R., Qian, W.-M., & Zhou, Z.-G. (2016b). Weighted composite quantile regression for single-index models. *J. Multivar. Anal.*, *148*, 34–48.
- Kai, B., Li, R., & Zou, H. (2010). Local composite quantile regression smoothing: an efficient and safe alternative to local polynomial regression. *J. R. Stat. Soc., B: Stat.*, *72*, 49–69.
- Kai, B., Li, R., & Zou, H. (2011). New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models. *Ann. Stat.*, *39*, 305 – 332.

- Keefe, M. J., Ferreira, M. A. R., & Franck, C. T. (2018). On the formal specification of sum-zero constrained intrinsic conditional autoregressive models. *Spat. Stat.*, *24*, 54–65.
- Koenker, R. (1984). A note on l-estimates for linear models. *Stat. Probab. Lett.*, *2*, 323–325.
- Koenker, R. (2005). *Quantile Regression*. Econometric Society Monographs. Cambridge University Press.
- Koenker, R., & Bassett, J. G. (1982). Robust tests for heteroscedasticity based on regression quantiles. *Econometrica*, (pp. 43–61).
- Lindley, D. V., & Smith, A. F. (1972). Bayes estimates for the linear model. *J. R. Stat. Soc., B: Stat. Methodol.*, *34*, 1–18.
- Luo, S., Zhang, C.-y., & Wang, M. (2019). Composite quantile regression for varying coefficient models with response data missing at random. *Symmetry*, *11*, 1065.
- Ma, X., & Zhang, J. (2016). Robust model-free feature screening via quantile correlation. *J. Multivar. Anal.*, *143*, 472–480.
- Ma, Y., & Yin, G. (2011). Censored quantile regression with covariate measurement errors. *Stat. Sin.*, *21*, 949.
- Neelon, B., Li, F., Burgette, L. F., & Benjamin Neelon, S. E. (2015). A spatiotemporal quantile regression model for emergency department expenditures. *Statist. Med.*, *34*, 2559–2575.
- Rue, H., & Held, L. (2005). *Gaussian Markov random fields: theory and applications*. Number 104 in Monographs on statistics and applied probability. Boca Raton: Chapman & Hall/CRC.
- Sun, J., Gai, Y., & Lin, L. (2013). Weighted local linear composite quantile estimation for the case of general error distributions. *J. Stat. Plan. Inference*, *143*, 1049–1063.
- Tian, Y., Lian, H., & Tian, M. (2017). Bayesian composite quantile regression for linear mixed effects models. *Commun. Stat. - Theory Methods*, *46*, 7717–7731.
- Tian, Y., Wang, L., Tang, M., & Tian, M. (2021). Weighted composite quantile regression for longitudinal mixed effects models with application to aids studies. *Commun. Stat. - Simul. Comput.*, *50*, 1837–1853.
- Xu, K. (2017). Model-free feature screening via a modified composite quantile correlation. *J. Stat. Plan. Inference*, *188*, 22–35.

- Zhao, W., Lian, H., & Song, X. (2017). Composite quantile regression for correlated data. *Comput. Stat. Data Anal.*, *109*, 15–33.
- Zhao, W.-h., Zhang, R.-q., Lü, Y.-z., & Liu, J.-c. (2016). Bayesian regularized regression based on composite quantile method. *Acta Math. Appl. Sin. Engl. Ser.*, *32*, 495–512.
- Zou, H., & Yuan, M. (2008). Composite quantile regression and the oracle model selection theory. *Ann. Stat.*, *36*, 1108–1126.