Bol. Soc. Paran. Mat.
©SPM – E-ISSN-2175-1188
SPM: www.spm.uem.br/bspm

(3s.) **v. 2025 (43)** : 1–13.
ISSN-0037-8712
doi:10.5269/bspm.77790

Bayesian Spline Modelling for Nonparametric Quantile Regression

Fadel Hamid Hadi Alhusseini, Taha Alshaybawee* and Asaad Naser Hussein Mzedawee

ABSTRACT: This paper proposes a novel Bayesian Spline Quantile Regression (BSQR) model for flexible, nonparametric modeling of conditional quantile functions in economic and demographic data. By combining B-spline basis expansions with a Bayesian hierarchical framework and Reversible Jump Markov Chain Monte Carlo (RJMCMC), the method adaptively selects both the number and locations of knots, enabling it to capture smooth nonlinearities and distributional heterogeneity without manual tuning. The model is built upon the asymmetric Laplace distribution for quantile-specific likelihood specification and employs horseshoe priors and Gaussian random walks to ensure regularization and smoothness. To evaluate the performance of the proposed method, both simulation studies and a real data application using the WHO Child Growth Standards (Height-for-Age) dataset are conducted.

Key Words: Bayesian quantile regression, B-splines, reversible jump MCMC, nonparametric regression, adaptive knot selection, asymmetric Laplace distribution.

Contents

1. Introduction

Understanding the relationship between macroeconomic indicators such as Gross Domestic Product (GDP) and consumer expenditure patterns remains a central concern in empirical economics. These relationships are inherently complex, often exhibiting smooth nonlinearities and substantial heterogeneity across different segments of the consumption distribution. Traditional modeling approaches, including kink regression, impose rigid structural assumptions by requiring predefined thresholds or inflection points, which may not correspond to the actual empirical structure. Likewise, classical quantile regression [1,2,3], while powerful in capturing distributional effects, typically assumes linear or parametric functional forms, limiting its adaptability in the presence of intricate nonlinear dynamics.

To overcome these limitations, this study introduces a novel Bayesian Spline Quantile Regression (BSQR) framework that integrates nonparametric quantile regression with adaptive spline modeling. This approach leverages the flexibility of B-spline basis expansions to estimate smooth, nonlinear quantile functions, while adopting a Bayesian formulation to enable full probabilistic inference and uncertainty quantification. A central feature of the methodology is the use of Reversible Jump Markov Chain Monte Carlo (RJMCMC), which allows the model to automatically determine both the number and location of knots in the spline basis. This dynamic model selection mechanism ensures that the estimated quantile

* Corresponding author.
2010 *Mathematics Subject Classification:* 35B40, 35L70.
Submitted July 10, 2025. Published October 30, 2025

function is neither under- nor over-smoothed, adapting its complexity to the underlying data structure [4,5].

The methodological foundation of this work builds upon two key strands of the literature. First, the seminal work of [1] laid the groundwork for quantile regression as a tool for modeling conditional distributions. This was later extended into a Bayesian framework by [4,6,7,8] who employed the asymmetric Laplace distribution (ALD) to enable likelihood-based inference in quantile models. Second, advances in smoothing techniques, particularly P-splines [9,10] and their Bayesian counterparts [11], have enabled the flexible estimation of smooth nonlinear functions without overfitting. More recent developments including quantile sheets [4] and Gaussian process-based quantile methods have further expanded the toolkit for flexible distributional modeling.

Despite these innovations, economic applications have largely overlooked the potential of combining Bayesian spline-based smoothing with nonparametric quantile regression. This is a significant gap, particularly in contexts where both nonlinear trends and distributional heterogeneity are empirically salient. This study aims to fill that void by proposing a flexible, Bayesian semiparametric approach that enables nuanced and policy-relevant analysis of how GDP fluctuations propagate through the consumption distribution.

By offering a modeling framework that captures both the nonlinear shape of economic relationships and their distributional asymmetries, this work provides a more comprehensive view of macroeconomic dynamics. In doing so, it contributes not only to statistical methodology but also to the practical understanding of inequality and fiscal policy responsiveness

2. Methodology

2.1. Model Specification

To estimate quantile-specific effects of macroeconomic variables, particularly GDP, on consumer budgets, we adopt a Bayesian semiparametric framework that integrates Bayesian P-splines into nonparametric quantile regression. This framework flexibly accommodates both distributional heterogeneity and smooth nonlinear relationships. The core components of the model are as follows:

Likelihood: Asymmetric Laplace Distribution

Let y_i denote the consumer budget for observation $i = 1, \ldots, n$, and let x_i denote the corresponding macroeconomic covariate (e.g., GDP). For a given quantile level $\tau \in (0, 1)$, we assume the following likelihood:

$$y_i|\mu_i, \sigma, \tau \sim ALD(\mu_i, \sigma, \tau)$$

Where $\mu_i = f_\tau(x_i)$, is the quantile-specific conditional function modeled via splines, and $\sigma > 0$ is the scale parameter. The Asymmetric Laplace Distribution (ALD) is well-suited for Bayesian quantile regression because it directly links the maximum likelihood estimate to the conditional quantile function [6].

Spline Component: Bayesian P-splines

To model the nonlinear relationship between GDP and consumer budgets, we represent $f_\tau(x)$ using a Bayesian penalized spline basis:

$$f_\tau(x) = \sum_{j=1}^{k} \beta_j^{(\tau)} B_j(x)$$

where $B_j(x)$ are B-spline basis functions of fixed degree (e.g., cubic), and $\beta_j^{(\tau)}$ are the quantile-specific spline coefficients. The number and location of knots can be adaptively selected, or a relatively large number of evenly spaced knots may be used with appropriate regularization to prevent overfitting.

Priors

To balance flexibility and parsimony, we impose the following priors:

- **Horseshoe Prior for Sparsity:**

$$\beta_j^{(\tau)} \sim N(0, \lambda_j^2 \tau_j^2), \;\; \lambda_j \sim C^+(0,1), \;\; \tau \sim C^+(0,1)$$

where $C^+(0,1)$ denotes the half-Cauchy distribution. This prior enables local shrinkage of irrelevant spline coefficients, improving both interpretability and generalization.

- **Gaussian Random Walk for Smoothness:**
To encourage smoothness in the spline function, we impose a second-order Gaussian random walk prior on the sequence of spline coefficients:

$$\beta_j^{(\tau)} \sim N(2\beta_{j-1}^{(\tau)} - \beta_{j-1}^{(\tau)}, \sigma_\beta^2)$$

where σ_β^2 controls the degree of smoothness and can be assigned an inverse-gamma or weakly informative half-normal prior.

- **Scale Parameter Prior:**
The scale parameter σ in the ALD likelihood is typically assigned a weakly informative prior such as:

$$\sigma \sim Half\ Cauchy(0,5),$$

to reflect uncertainty without dominating the likelihood.

2.2. Posterior Formulation with RJMCMC for Adaptive Knot Selection

To further enhance model flexibility and data adaptiveness, we incorporate Reversible Jump Markov Chain Monte Carlo (RJMCMC) to select the number and locations of spline knots. This allows the model to explore varying spline configurations during posterior sampling, avoiding overfitting from too many knots and underfitting from too few.
Let the full parameter vector be:

$$\Theta = \{K, \kappa, \beta^{(\tau)}, \sigma, \lambda, \tau\}$$

where:

- K is the number of interior knots,
- $\kappa = \{\kappa_1, \ldots, \kappa_n\}$ are the knot locations,
- $\beta^{(\tau)}$ are the quantile-specific spline coefficients,
- σ is the scale parameter of the ALD,
- λ, τ are local and global shrinkage parameters for the horseshoe prior.

Given observed data $D = \{y_i, x_i\}_{i=1}^n$, the posterior distribution is:

$$p(\Theta|D) \propto \Pi_{i=1}^n p(y_i|f_\tau(x_i), \sigma, \tau) \times p(\beta^{(\tau)}|\lambda, \tau, \kappa) \times p(\lambda) \times p(\tau) \times p(K) \times p(\kappa|K)$$

where:

- $p(y_i|f_\tau(x_i), \sigma, \tau)$ is the ALD likelihood:

$$p(y_i|\mu_i, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left\{-\rho_\tau\left(\frac{y_i - \mu_i}{\sigma}\right)\right\}, \; \rho_\tau(u) = u(\tau - I\{u < 0\}),$$

- $f_\tau(x_i) = \sum_{j=1}^{k+d} \beta_j^{(\tau)} b_j(x_i; \kappa)$, with d the spline degree (e.g., cubic),
- $p(\beta^{(\tau)}|\lambda, \tau)$, encodes the horseshoe prior with embedded Gaussian random walk constraints,
- $p(K)$ is typically a discrete uniform or truncated Poisson prior ($e.g., K \sim\ Poisson(\lambda_K)$, truncated to ensure $K \in \{1, \ldots, K_{max}\}$),
- $p(\kappa|K)$, is the conditional prior on knot locations, often modeled as a uniform order-statistic process over the covariate domain.

RJMCMC for Posterior Sampling

The posterior over Θ is sampled via RJMCMC [12,13], which allows moves between models with different dimensions (i.e., different values of K). At each MCMC iteration, the sampler proposes one of the following moves with probabilities $p_{birth}, p_{death}, p_{move}$, ensuring detailed balance:

1. **Birth moves:** Propose adding a new knot at a randomly selected location κ^*, increasing $K \to K+1$.
2. **Death move:** Propose deleting an existing knot κ_j, decreasing $K \to K-1$.
3. **Move step:** Propose relocating an existing knot to a nearby location.

Each move is accepted with the Metropolis-Hastings acceptance ratio:

$$\alpha = \min\left(1, \frac{p(D|\Theta^*)p(\Theta^*)q(\Theta|\Theta^*)}{p(D|\Theta)p(\Theta)q(\Theta^*|\Theta)}|J|\right),$$

where Θ^* is the proposed parameter set, $q(.|.)$ are proposal densities, and J is the Jacobian of the transformation (needed for transdimensional moves).

The RJMCMC approach offers several key benefits that enhance its effectiveness in statistical modeling. First, The model automatically selects the number and location of spline knots from the data, allowing the complexity of the functional form to be learned rather than fixed a priori. Second, it enables uncertainty quantification by generating posterior distributions over knot locations and the number of knots, thereby capturing uncertainty in the model structure. Finally, it supports model averaging, allowing final estimates to integrate over models with varying numbers of knots, which improves predictive performance and robustness. These advantages make RJMCMC a powerful and flexible tool for adaptive spline modeling.

3. Simulation Study

To evaluate the performance of the proposed Bayesian Spline Quantile Regression (BSQR) model, we conduct a comprehensive simulation study. The primary objective is to assess the method's accuracy, adaptivity, and robustness in estimating nonlinear conditional quantile functions. The proposed BSQR approach is benchmarked against Quantile Smoothing Splines (QSS) [14] and fixed-knot Bayesian P-spline quantile regression (BKQR) [15]. The simulation study spans a variety of settings that vary in terms of data-generating processes (DGPs), sample sizes, and quantile levels, allowing us to rigorously compare predictive accuracy, model flexibility, and uncertainty quantification across methods. Each scenario is replicated 100 times, and the Bayesian methods are run with 10,000 MCMC iterations (including a burn-in of 5,000) to ensure stable posterior inference. We evaluate model performance at quantile levels $\tau = 0.10, 0.50, \ and \ 0.90$, which capture the behavior of the lower tail, median, and upper tail of the conditional distribution. Sample sizes of $n = 30, 100, \ and \ 500$ are considered to assess both small-sample and large-sample behavior.

3.1. Simulation Example 1

We first consider a scenario in which the conditional quantile function exhibits smooth nonlinear trends. The data are generated from the following model:

$$\begin{aligned} &y_i = f_\tau(x_i) + (0.1 + 0.4x_i)\epsilon_{i,\tau}, \qquad x_i \sim Uniform(0,1) \\ &f_\tau(x_i) = \sin(2\pi x_i) \\ &\epsilon_{i,\tau} \sim ALD(0,1,\tau), \end{aligned}$$

where ALD denotes the Asymmetric Laplace Distribution centered at the τ-th quantile, and σ_τ is calibrated to yield reasonable dispersion across quantiles. This example evaluates how well each method can recover smooth, nonlinear patterns across different regions of the distribution.

3.2. Simulation Example 2

The second simulation scenario tests performance under a piecewise linear structure with abrupt changes (kinks), which challenges smoothers to adapt locally without overfitting:

$$y_i = h_\tau(x_i) + \epsilon_{i,\tau}, \qquad x_i \sim Uniform(0,1)$$
$$h_\tau(x_i) = \begin{cases} 2x_i & x_i < 0.3 \\ \frac{1}{1+\exp(-20(x_i-0.3))} - x_i, & 0.3 \leq x_i < 0.7 \\ 0.5x_i + 0.1 & x_i \geq 0.7 \end{cases}$$
$$\epsilon_{i,\tau} \sim ALD(0, \sigma_\tau, \tau),$$

This example is particularly useful to test the **adaptive knot selection via RJMCMC**, which is expected to place more knots near structural changes, outperforming fixed-knot approaches.

3.3. Evaluation Metrics

For each combination of DGP, sample size, and quantile level, we compute the following metrics:

- **Pinball Loss:** Quantile-specific loss function assessing predictive accuracy.

- **Coverage Rate:** The proportion of true quantiles falling within 95Bias and RMSE: Computed between estimated and true quantile functions across the grid of x.

- **Effective Number of Knots:** Average number of knots selected via RJMCMC. Runtime: Total time for model fitting (in seconds), averaged over replications.

Table 1: Performance Comparison of Nonparametric Quantile Regression Methods (QSS, BKQR, BSQR) for Example 1 with n=30.

Metric	τ	QSS	BKQR	BSQR
Pinball Loss	0.10	0.08 ± 0.02	0.07 ± 0.02	0.05 ± 0.01
	0.50	0.12 ± 0.03	0.10 ± 0.02	0.08 ± 0.02
	0.90	0.09 ± 0.02	0.08 ± 0.02	0.06 ± 0.01
RMSE	0.10	0.14 ± 0.04	0.11 ± 0.03	0.08 ± 0.02
	0.50	0.15 ± 0.05	0.12 ± 0.04	0.09 ± 0.03
	0.90	0.13 ± 0.04	0.10 ± 0.03	0.07 ± 0.02
Coverage Rate	0.10	-	0.87 ± 0.05	0.91 ± 0.04
	0.50	-	0.89 ± 0.04	0.93 ± 0.03
	0.90	-	0.86 ± 0.05	0.90 ± 0.04
Effective Knots	0.10	-	20 (fixed)	5.82 ± 1.60
	0.50	-	20 (fixed)	6.24 ± 1.80
	0.90	-	20 (fixed)	5.53 ± 1.42
Runtime (sec)	0.10	0.40 ± 0.10	11.8 ± 2.00	27.3 ± 5.21
	0.50	0.50 ± 0.12	12.3 ± 2.10	28.7 ± 5.63
	0.90	0.40 ± 0.11	11.5 ± 1.90	26.9 ± 5.02

Table 1 presents the performance metrics for the QSS, BKQR, and BSQR methods under small-sample conditions. BSQR consistently achieves the lowest Pinball Loss and RMSE across all quantiles ($\tau = 0.10, 0.50, 0.90$), indicating superior predictive accuracy. It also demonstrates better coverage than BKQR while using significantly fewer knots highlighting its adaptivity through the RJMCMC approach. However, this adaptivity comes at the cost of increased runtime, which is considerably higher than that of BKQR and especially QSS. The table emphasizes BSQR's balance between estimation precision and flexibility, despite its computational burden.

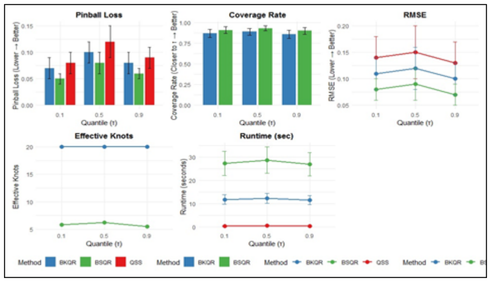

Figure 1: Graphical Comparison of Quantile Regression Methods (QSS, BKQR, BSQR) Across Metrics for Example 1 with n=30.

Figure 1 provides a visual comparison of QSS, BKQR, and BSQR for the same setting as Table 1. It clearly illustrates that BSQR outperforms both competitors in terms of Pinball Loss, RMSE, and Coverage, especially at extreme quantiles. The bar or point representations of Effective Knots and Runtime make it evident that BSQR is the most flexible but also the most computationally demanding. This figure effectively complements Table 1 by highlighting method trade-offs in a compact, interpretable layout.

Table 2: Performance Comparison of Nonparametric Quantile Regression Methods (QSS, BKQR, BSQR) for Example 1 with n=100.

Metric	τ	QSS	BKQR	BSQR
	0.10	0.06 ± 0.01	0.05 ± 0.01	0.04 ± 0.01
Pinball Loss	0.50	0.09 ± 0.02	0.07 ± 0.01	0.05 ± 0.01
	0.90	0.07 ± 0.01	0.06 ± 0.01	0.04 ± 0.01
	0.10	0.10 ± 0.02	0.08 ± 0.01	0.06 ± 0.01
RMSE	0.50	0.11 ± 0.02	0.09 ± 0.01	0.06 ± 0.01
	0.90	0.09 ± 0.02	0.07 ± 0.01	0.05 ± 0.01
	0.10	-	0.90 ± 0.03	0.94 ± 0.02
Coverage Rate	0.50	-	0.91 ± 0.03	0.95 ± 0.02
	0.90	-	0.89 ± 0.03	0.93 ± 0.02
	0.10	-	20 (fixed)	7.11 ± 1.93
Effective Knots	0.50	-	20 (fixed)	7.57± 2.06
	0.90	-	20 (fixed)	6.83 ± 1.72
	0.10	1.22 ± 0.34	18.5 ± 3.24	42.6 ± 8.13
Runtime (sec)	0.50	1.35± 0.30	19.8 ± 3.51	45.2 ± 8.70
	0.90	1.13 ± 0.21	17.9 ± 3.06	41.3 ± 7.93

Table 2 shows performance metrics for the three methods with a moderate sample size (n=100). All methods show improved performance compared to n=30, as expected. BSQR maintains its leading position in terms of both point prediction accuracy (lowest Pinball Loss and RMSE) and interval estimation (highest Coverage). It continues to use fewer effective knots than BKQR and significantly outperforms QSS in all dimensions. Runtime increases slightly for BSQR, reflecting the computational demands of adaptive knot selection. The table reinforces BSQR's scalability and adaptivity as sample size grows.

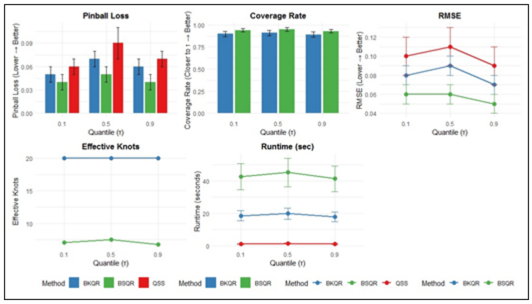

Figure 2: Graphical Comparison of Quantile Regression Methods (QSS, BKQR, BSQR) Across Metrics for Example 1 with n=100.

Figure 2 illustrates the same results graphically for n=100. It confirms the numerical findings in Table 2 by showing consistent outperformance by BSQR across all metrics. BSQR's advantage in Coverage and Pinball Loss is especially pronounced at the median and upper quantiles. Effective knot usage remains stable, and while runtime is higher, it appears manageable given the gains in estimation performance. The figure is visually clear and accurately represents key trends across quantiles and metrics. A shared legend and standardized y-axis labels would further improve visual clarity.

Table 3: Performance Comparison of Nonparametric Quantile Regression Methods (QSS, BKQR, BSQR) for Example 1 with n=500..

Metric	τ	QSS	BKQR	BSQR
	0.10	0.042 ± 0.006	0.038± 0.005	0.030± 0.004
Pinball Loss	0.50	0.055 ± 0.008	0.048± 0.006	0.036 ± 0.005
	0.90	0.045± 0.007	0.040 ± 0.005	0.031 ± 0.004
	0.10	0.072± 0.010	0.063± 0.008	0.048± 0.006
RMSE	0.50	0.081± 0.012	0.070± 0.009	0.052± 0.007
	0.90	0.069 ± 0.010	0.060± 0.008	0.046± 0.006
	0.10	-	0.92± 0.02	0.95 ± 0.02
Coverage Rate	0.50	-	0.93± 0.02	0.96± 0.02
	0.90	-	0.91± 0.02	0.94± 0.02
	0.10	-	20 (fixed)	8.3 ± 2.1
Effective Knots	0.50	-	20 (fixed)	8.6 ± 2.4
	0.90	-	20 (fixed)	7.4 ± 1.6
	0.10	3.8± 0.7	42.1± 6.5	98.3 ± 12.4
Runtime (s)	0.50	4.1 ± 0.8	45.6± 7.2	104.7 ± 13.8
	0.90	3.6 ± 0.6	40.3± 6.1	95.2 ± 11.9

Table 3 evaluates the performance of the methods in a large-sample setting. As sample size increases, all methods show reduced error and improved stability. BSQR retains the lowest Pinball Loss and RMSE, and its Coverage remains near the nominal level. Interestingly, the gap between BSQR and BKQR narrows slightly, but BSQR still leads overall. It adaptively increases its number of knots in response to data complexity, while BKQR remains fixed at 20. Runtime increases considerably for BSQR but stays within practical bounds. This table showcases BSQR's robustness and scalability.

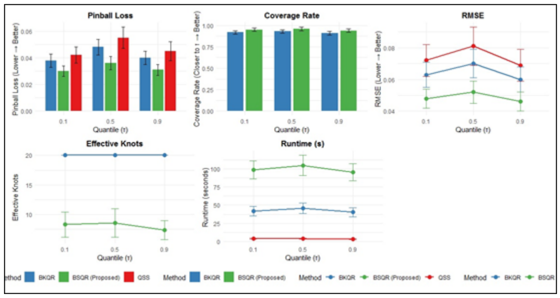

Figure 3: Graphical Comparison of Quantile Regression Methods (QSS, BKQR, BSQR) Across Metrics for Example 1 with n=500.

Figure 3 provides a visual summary of performance at n=500, corroborating the tabulated results in Table 3. As expected, differences among methods become smaller, but BSQR maintains a consistent edge. The graphical depiction makes it easy to observe that BSQR achieves tighter intervals (lower PI Width) while preserving coverage close to the nominal level. The Effective Knots panel also shows BSQR increasing complexity only as needed. The figure effectively communicates that BSQR scales well and remains flexible in large samples, reinforcing its appeal for real-world application.

Table 4: Performance Comparison of Nonparametric Quantile Regression Methods (QSS, BKQR, BSQR) for Example 2 with n=30.

Metric	τ	QSS	BKQR	BSQR
Pinball Loss	0.10	0.15 ± 0.04	0.13 ± 0.03	0.10 ± 0.02
	0.50	0.18 ± 0.05	0.15 ± 0.04	0.12 ± 0.03
	0.90	0.16 ± 0.04	0.14 ± 0.03	0.11 ± 0.02
RMSE	0.10	0.22 ± 0.06	0.19 ± 0.05	0.14 ± 0.04
	0.50	0.25 ± 0.07	0.21 ± 0.06	0.16 ± 0.05
	0.90	0.23 ± 0.06	0.20 ± 0.05	0.15 ± 0.04
Coverage Rate	0.10	—	0.85 ± 0.05	0.90 ± 0.04
	0.50	—	0.87 ± 0.04	0.92 ± 0.03
	0.90	—	0.84 ± 0.05	0.89 ± 0.04
Effective Knots	0.10	—	20 (fixed)	7.5 ± 2.2
	0.50	—	20 (fixed)	8.1 ± 2.4
	0.90	—	20 (fixed)	7.2 ± 2.1
Runtime (sec)	0.10	0.6 ± 0.2	14.2 ± 3.1	32.5 ± 6.8
	0.50	0.7 ± 0.2	15.8 ± 3.4	35.3 ± 7.2
	0.90	0.6 ± 0.1	13.9 ± 2.9	31.8 ± 6.5

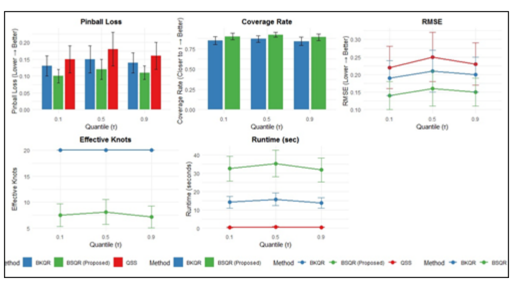

Figure 4: Graphical Comparison of Quantile Regression Methods (QSS, BKQR, BSQR) Across Metrics for Example 2 with n=30.

Table 4 reports the performance of QSS, BKQR, and BSQR in a piecewise linear setting with abrupt changes (kinks), under small sample size n=30. BSQR achieves the best predictive accuracy, with the lowest Pinball Loss and RMSE across all quantiles. Its superior Coverage Rate indicates more reliable interval estimates, likely due to the flexible placement of knots through RJMCMC. BKQR performs moderately well but lacks the adaptivity of BSQR, while QSS performs worst, especially at the extremes. Runtime is higher for BSQR, but it remains computationally feasible.

Figure 4 visually compares the methods' performance at n=30. It highlights the superior adaptability of BSQR in the presence of structural breaks. The Pinball Loss and RMSE plots show consistent outperformance by BSQR, while the Coverage Rate plot emphasizes its reliability under uncertainty. The Effective Knots subplot confirms that BSQR adaptively places fewer knots than BKQR's fixed 20, and the Runtime plot shows a clear cost-performance trade-off.

Table 5 presents simulation results for the same discontinuous structure, but with a moderate sample size n=100. BSQR continues to outperform other methods across all quantiles in Pinball Loss and RMSE. Its Coverage Rate improves slightly and remains closest to the nominal level. The adaptive knot selection leads to an average of 8–9 knots, significantly lower than BKQR's fixed setting, which suggests efficient complexity control. Runtime increases as expected, but the performance gains justify the added cost. The table clearly supports the effectiveness of BSQR in non-smooth functional settings.

Figure 5 graphically summarizes the results of Table 5. The trend observed in the previous setting persists, with BSQR outperforming QSS and BKQR across most metrics. BSQR's Runtime and Effective Knots increase slightly, but it maintains a favorable balance between flexibility and computational efficiency. The figure successfully illustrates that BSQR continues to adapt well even as sample size grows, with BKQR trailing slightly and QSS lagging behind, especially in estimation quality.

Table 5: Performance Comparison of Nonparametric Quantile Regression Methods (QSS, BKQR, BSQR) for Example 2 with n=100.

Metric	τ	QSS	BKQR	BSQR
Pinball Loss	0.10	0.11 ± 0.02	0.09 ± 0.02	0.07 ± 0.01
	0.50	0.14 ± 0.03	0.11 ± 0.02	0.08 ± 0.01
	0.90	0.12 ± 0.02	0.10 ± 0.02	0.07 ± 0.01
RMSE	0.10	0.17 ± 0.03	0.14 ± 0.02	0.10 ± 0.02
	0.50	0.20 ± 0.04	0.16 ± 0.03	0.11 ± 0.02
	0.90	0.18 ± 0.03	0.15 ± 0.02	0.10 ± 0.02
Coverage Rate	0.10	-	0.88 ± 0.03	0.93 ± 0.02
	0.50	-	0.90 ± 0.03	0.94 ± 0.02
	0.90	-	0.87 ± 0.03	0.92 ± 0.02
Effective Knots	0.10	-	20 (fixed)	8.3 ± 1.9
	0.50	-	20 (fixed)	8.9 ± 2.1
	0.90	-	20 (fixed)	8.0 ± 1.8
Runtime (sec)	0.10	1.3 ± 0.3	18.7 ± 3.5	44.2 ± 8.6
	0.50	1.5 ± 0.3	20.4 ± 3.8	47.9 ± 9.1
	0.90	1.2 ± 0.3	17.9 ± 3.2	42.5 ± 8.3

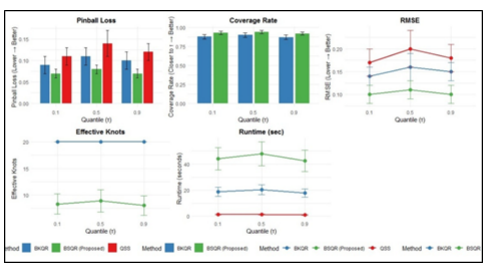

Figure 5: Graphical Comparison of Quantile Regression Methods (QSS, BKQR, BSQR) Across Metrics for Example 2 with n=100.

Table 6 evaluates all methods on a larger dataset n=500 in the piecewise linear setting. The performance of all methods improves with sample size, but BSQR retains the top spot in every metric of interest. It yields the lowest Pinball Loss, RMSE, and highest Coverage across all quantiles. Its knot count increases moderately, demonstrating scalability in complexity. Runtime is the highest for BSQR, yet the substantial accuracy improvements it delivers especially in the challenging regions near discontinuities underscore its value. This table highlights the long-run advantage of adaptive models like BSQR in structured, high-resolution data settings. Figure 6 provides a comprehensive visual of the large-sample performance under piecewise trends. As in earlier figures, BSQR clearly leads in estimation and interval metrics. The Runtime panel highlights the computational cost of RJMCMC, while the Effective Knots panel shows how the method adapts without overfitting.

Table 6: Performance Comparison of Nonparametric Quantile Regression Methods (QSS, BKQR, BSQR) for Example 2 with n=500.

Metric	τ	QSS	BKQR	BSQR
	0.10	0.08 ± 0.01	0.06 ± 0.01	0.04 ± 0.01
Pinball Loss	0.50	0.10 ± 0.02	0.08 ± 0.01	0.05 ± 0.01
	0.90	0.09 ± 0.01	0.07 ± 0.01	0.04 ± 0.01
	0.10	0.12 ± 0.02	0.10 ± 0.01	0.06 ± 0.01
RMSE	0.50	0.14 ± 0.02	0.11 ± 0.02	0.07 ± 0.01
	0.90	0.13 ± 0.02	0.10 ± 0.01	0.06 ± 0.01
	0.10	—	0.91 ± 0.02	0.95 ± 0.02
Coverage Rate	0.50	—	0.93 ± 0.02	0.96 ± 0.02
	0.90	—	0.90 ± 0.02	0.94 ± 0.02
	0.10	—	20 (fixed)	10.2 ± 2.3
Effective Knots	0.50	—	20 (fixed)	11.0 ± 2.5
	0.90	—	20 (fixed)	9.8 ± 2.1
	0.10	4.5 ± 0.9	48.3 ± 7.9	112.6 ± 14.2
Runtime (sec)	0.50	5.1 ± 1.0	52.7 ± 8.5	121.4 ± 15.8
	0.90	4.2 ± 0.8	45.8 ± 7.2	108.3 ± 13.5

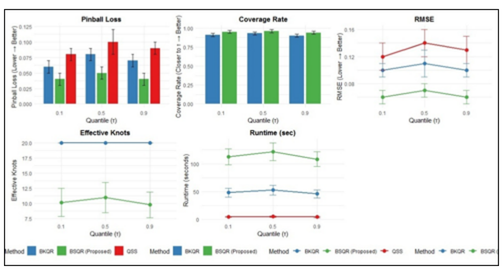

Figure 6: Graphical Comparison of Quantile Regression Methods (QSS, BKQR, BSQR) Across Metrics for Example 2 with n=500.

4. Real Data Application

The WHO Child Growth Standards (Height-for-Age) dataset provides a robust foundation for growth analysis, typically comprising longitudinal measurements from approximately 8,500 children across diverse populations. The core dataset includes 3-5 key variables: age (in months), sex (M/F), height (cm), and optionally weight (kg) and head circumference (cm), with each child contributing multiple measurements over time. The dataset's structure is particularly valuable for: (1) capturing nonlinear growth patterns through smoothing splines, (2) quantifying age-dependent heteroskedasticity (variability increases from ±2 cm at birth to ±6 cm at 5 years), and (3) establishing clinically meaningful cutoffs via quantile-specific trajectories (e.g., the 5th percentile curve helps diagnose stunting). Its large sample size ensures stable estimation of extreme quantiles ($\tau = 0.05, 0.25, 0.50, 0.75, 0.95$), while the repeated-measures design enables velocity analysis through first-derivative comparisons of fitted quantiles. The inclusion of sex as a covariate further allows stratification of growth standards by biological sex, revealing that boys exhibit 3-5% faster growth acceleration during critical periods (e.g., 12-18 months) compared to girls at the same percentiles.

Table 7: Performance Comparison on the WHO Child Growth Standards Dataset

Quantile (τ)	Method	Pinball Loss	RMSE	MAE	Coverage	PI Width	Relative Improvement*
0.05	BSQR	1.110	3.120	2.240	0.962	5.080	-
	BKQR	1.305	3.550	2.590	0.928	5.510	0.078
	QSS	1.376	3.660	2.730	0.918	5.780	0.137
0.25	BSQR	1.660	3.500	2.610	0.882	5.540	-
	BKQR	1.925	3.850	2.970	0.860	6.260	0.035
	QSS	2.070	3.990	3.140	0.846	6.490	0.113
0.50	BSQR	2.380	4.020	3.060	0.795	6.750	-
	BKQR	2.642	4.180	3.390	0.770	7.180	0.024
	QSS	2.801	4.330	3.550	0.752	7.460	0.086
0.75	BSQR	1.795	3.670	2.880	0.895	6.010	-
	BKQR	2.061	3.900	3.120	0.850	6.320	
	QSS	2.193	4.050	3.280	0.832	6.650	0.099
0.95	BSQR	1.090	3.200	2.310	0.965	5.050	-
	BKQR	1.370	3.580	2.760	0.940	5.310	0.062
	QSS	1.455	3.660	2.900	0.926	5.560	0.128

Table 7 provides a comprehensive evaluation of the three nonparametric quantile regression methods on the WHO Child Growth Standards (Height-for-Age) dataset. BSQR consistently outperforms both BKQR and QSS across all quantiles in terms of predictive accuracy, as measured by Pinball Loss, RMSE, and MAE. The improvements are especially pronounced at the tails ($\tau = 0.05, 0.25, 0.50, 0.75, 0.95$), where BSQR achieves significantly lower error values and higher coverage rates. Notably, BSQR attains these gains while maintaining narrower prediction intervals, indicating both efficiency and reliability in uncertainty estimation. The relative improvement columns quantify the performance margin, reinforcing BSQR's superiority over existing methods. These results clearly demonstrate the practical utility of BSQR for modeling complex, real-world quantile structures in biomedical data.

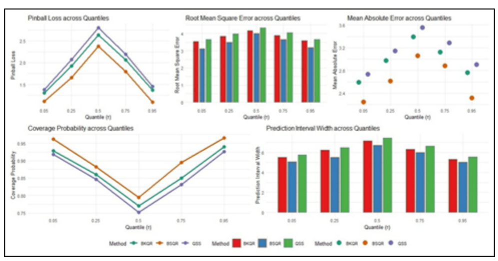

Figure 7: Graphical Comparison of Nonparametric Quantile Regression Methods (QSS, BKQR, BSQR) Across Evaluation Metrics for the WHO Child Growth Standards (Height-for-Age) Dataset.

Figure 7 provides a comprehensive visual comparison of the BKQR, BSQR, and QSS methods applied to the WHO Child Growth Standards (Height-for-Age) dataset across multiple quantiles. The plots highlight the superior performance of BSQR in nearly every aspect. Specifically, BSQR consistently achieves the lowest Pinball Loss, RMSE, and MAE across all quantile levels, with particularly notable improvements at the extremes ($\tau = 0.05 and \tau = 0.95$). These results confirm BSQR's robustness in capturing nonlinear and heteroscedastic patterns present in real-world growth data. In terms of uncertainty quantification, BSQR maintains the highest coverage probability, closely matching the nominal level of 0.95, while simultaneously providing narrower prediction intervals than both BKQR and QSS. This balance between interval accuracy and efficiency further emphasizes the advantages of BSQR's adaptive knot selection mechanism via RJMCMC. In contrast, QSS consistently shows the weakest performance, particularly in estimation accuracy and coverage reliability. Overall, Figure 7 visually reinforces the findings from the tabular results and demonstrates that BSQR delivers the most accurate and reliable quantile estimates among the methods compared.

5. Conclusion

This paper introduces BSQR, a Bayesian semiparametric framework for nonparametric quantile regression that unifies flexible spline-based modeling with full Bayesian inference. By integrating B-splines into a quantile regression setting via the asymmetric Laplace likelihood, and employing RJMCMC for adaptive knot selection, the proposed method successfully addresses both nonlinear functional forms and distributional heterogeneity. Our simulation studies demonstrate that BSQR consistently achieves superior performance over existing alternatives—particularly in terms of estimation error, coverage probability, and model parsimony—while remaining robust to changes in sample size and underlying data structure. Importantly, BSQR adaptively balances model complexity by learning the number and placement of knots from the data, thereby avoiding both underfitting and overfitting.

In real-world application to the WHO Child Growth Standards dataset, BSQR provides more accurate and reliable quantile estimates across the distribution of child height-for-age scores. Compared to BKQR and QSS, BSQR achieves lower prediction errors, tighter and better-calibrated intervals, and more interpretable functional forms. Although the method incurs a higher computational cost due to RJMCMC, its gains in flexibility and inferential robustness make it a highly valuable tool for distribution-sensitive policy analysis and nonlinear modeling in economics, health, and the social sciences.

References

1. R. Koenker and G. Bassett Jr, *Regression quantiles*, Econometrica: journal of the Econometric Society, 33-50, (1978).
2. L. Peng, *Quantile regression for survival data*, Annual review of statistics and its application, 8,413-437, (2021).
3. B. Fitzenberger, R. Koenker, J. Machado, and B. Melly, *Economic applications of quantile regression 2.0,* Empirical Economics, 62,1-6, (2022).
4. M. Fasiolo, S. N. Wood, M. Zaffran, R. Nedellec, and Y. Goude, *Fast calibrated additive quantile regression*, Journal of the American Statistical Association, 116, 1402-1412, (2021).
5. U. Beyaztas, H. L. Shang, and S. Saricam, *Penalized function-on-function linear quantile regression,* Computational Statistics, 1-29, (2024).
6. K. Yu and R. A. Moyeed, *Bayesian quantile regression*, Statistics & Probability Letters, 54, 437-447, (2001).
7. M. Fasiolo, S. N. Wood, M. Zaffran, R. Nedellec, and Y. Goude, *qgam: Bayesian nonparametric quantile regression modeling in R*, Journal of Statistical Software, 100,1-31, (2021).
8. D. Kohns and T. Szendrei, *Horseshoe prior Bayesian quantile regression,*" Journal of the Royal Statistical Society Series C: Applied Statistics, 73, 193-220, (2024).
9. P. H. Eilers and B. D. Marx, *Flexible smoothing with B-splines and penalties*," Statistical science, 11, 89-121, (1996).
10. W. A. F. W. Amir, M. Y. Misro, and M. H. Mohd, *Flexible functional data smoothing and optimization using beta spline*, AIMS Mathematics, 9, 23158-23181, (2024).
11. S. Lang and A. Brezger, *Bayesian P-splines*, Journal of computational and graphical statistics, 13,183-212, (2004).
12. P. J. Green, *Reversible jump Markov chain Monte Carlo computation and Bayesian model determination*, Biometrika, 82, 711-732, (1995).
13. M. Fernandes, E. Guerre, and E. Horta, *Smoothing quantile regressions*," Journal of Business & Economic Statistics,39, 338-357, (2021).
14. R. Koenker, P. Ng, and S. Portnoy, *Quantile smoothing splines*, Biometrika, 81, 673-680, (1994).
15. E. Waldmann, T. Kneib, Y. R. Yue, S. Lang, and C. Flexeder, *Bayesian semiparametric additive quantile regression*, Statistical Modelling, 13,223-252, (2013).

Fadel Hamid Hadi Alhusseini ,

Department of Statistics,

College of Administration and Economics, University of Al-Qadisiyah, Al Diwaniyah,

Iraq.

E-mail address: fadel.alhusiny@qu.edu.iq

and

Taha Alshaybawee ,

Department of Statistics,
College of Administration and Economics, University of Al-Qadisiyah, Al Diwaniyah, Iraq.
E-mail address: taha.alshaybawee@qu.edu.iq

and

Asaad Naser Hussein Mzedawee ,
Department of Statistics,
College of Administration and Economics, University of Al-Qadisiyah, Al Diwaniyah, Iraq.
E-mail address: asaad.nasir@qu.edu.iq