FLEXIBLE BAYESIAN QUANTILE REGRESSION IN ORDINAL MODELS

Mohammad Arshad Rahman[a] and
Shubham Karnawat[b]

[a]*Indian Institute of Technology Kanpur, India*
[b]*Credit Suisse, India*

ABSTRACT

This article is motivated by the lack of flexibility in Bayesian quantile regression for ordinal models where the error follows an asymmetric Laplace (AL) distribution. The inflexibility arises because the skewness of the distribution is completely specified when a quantile is chosen. To overcome this shortcoming, we derive the cumulative distribution function (and the moment-generating function) of the generalized asymmetric Laplace (GAL) distribution – a generalization of AL distribution that separates the skewness from the quantile parameter – and construct a working likelihood for the ordinal quantile model. The resulting framework is termed flexible Bayesian quantile regression for ordinal (FBQROR) models. However, its estimation is not straightforward. We address estimation issues and propose an efficient Markov chain Monte Carlo (MCMC) procedure based on Gibbs sampling and joint Metropolis–Hastings algorithm. The advantages of the proposed model are demonstrated in multiple simulation studies and implemented to analyze public opinion on homeownership as the best long-term investment in the United States following the Great Recession.

Keywords: Generalized asymmetric Laplace distribution; Gibbs sampling; Great Recession; homeownership; Markov chain Monte Carlo; Metropolis–Hastings

Topics in Identification, Limited Dependent Variables, Partial Observability, Experimentation, and Flexible Modeling
Advances in Econometrics, Volume 40B, 211–251

ISSN: 0731-9053/doi:10.1108/S0731-90532019000040B011

1. INTRODUCTION

Quantile regression, proposed by Koenker and Bassett (1978), models the conditional quantiles of the dependent variable as a function of the covariates. This method is particularly useful if interest lies in the outer regions of the conditional distribution and/or the data violates the standard assumptions of mean regression (e.g., the presence of heteroscedasticity, existence of outliers, and so on). Since its introduction, the concept has gained considerable attention from researchers worldwide and across ideologies. Within the frequentist econometrics/statistics literature, the advantages of quantile regression estimators are well studied and the computational challenges pertaining to optimizing a non-differentiable loss/objective function have been adequately dealt with. A valuable source is the study by Koenker (2005) and references therein. The development of Bayesian quantile regression faced an impediment since errors in quantile regression were not assumed to follow any distribution (necessary for writing the likelihood). About two decades later, Koenker and Machado (1999) noted that the quantile loss function appears in the exponent of an asymmetric Laplace (AL) distribution (Kotz, Kozubowski, & Podgorski, 2001; Yu & Zhang, 2005), thus facilitating the construction of a parametric likelihood. This distribution was utilized by Yu and Moyeed (2001) to propose a Bayesian method for estimating quantile regression in linear models. The estimation algorithm was further refined by Tsionas (2003), Reed and Yu (2009), and recently Kozumi and Kobayashi (2011) proposed a Gibbs sampling algorithm, where they exploit the normal-exponential mixture representation of the AL distribution. The AL likelihood has been utilized to develop algorithms for Bayesian quantile regression in Tobit models (Kozumi & Kobayashi, 2011; Yu & Stander, 2007), Tobit models with endogenous covariates (Kobayashi, 2017), censored models (Reich & Smith, 2013), censored dynamic panel data models (Kozumi & Kobayashi, 2012), count data models (Lee & Neocleous, 2010), and mixed-effect or longitudinal data models (Geraci & Bottai, 2007; Luo, Lian, & Tian, 2012).

Quantile regression in ordinal models is different since the dependent variable takes discrete and ordered values (which has no cardinal interpretation) and does not yield continuous quantiles. Ordinal outcomes typically arise as response to surveys, and applications are common in economics, finance, marketing, and the social sciences. Similar to the continuous case, interest in ordinal quantile regression is aimed to provide a much richer view of the heterogeneous effect of the covariates on the outcomes. However, estimation is more challenging. A frequentist approach using simulated annealing was proposed by Zhou (2010). Bayesian estimation of ordinal quantile regression was introduced by Rahman (2016) and extended to longitudinal data models by Alhamzawi and Ali (2018). A special case of ordinal model is the binary model, where the outcome variable is dichotomous (i.e., takes only two values, typically coded as 1 for "success" and 0 for "failure"). Bayesian quantile regression in binary models was proposed by Benoit and Poel (2010) and employed to study the mode of transportation to work. Rahman and Vossmeyer (2019) extended Bayesian quantile regression to

binary longitudinal outcomes and proposed an efficient Markov chain Monte Carlo (MCMC) algorithm for its estimation. The Bayesian ordinal quantile regression model has been utilized in a wide variety of studies including evaluation of credit risk (Miguéis, Benoit, & Poel, 2013), educational attainment (Rahman, 2016), public opinion on tax policy (Rahman, 2016), public opinion on nuclear power plants operation (Omata, Katayama, & Arimura, 2017), and illness severity (Alhamzawi & Ali, 2018).

The list of articles on Bayesian quantile regression mentioned above, although incomplete, clearly affirm that the AL distribution has played a crucial role in the development of Bayesian quantile regression. However, the AL distribution poses a critical limitation since a single parameter defines both the quantile and the skewness of the distribution. In addition, the mode of the distribution is always fixed at the location parameter value for all quantiles. To overcome these drawbacks, Yan and Kottas (2017) proposed the probability density function (*pdf*) of the generalized asymmetric Laplace (GAL) distribution by introducing a shape parameter into the mean of the normal kernel in the AL mixture representation. The GAL distribution uses different parameters for quantile and skewness and thus adds much-needed flexibility for Bayesian quantile regression. They utilized the GAL *pdf* and proposed algorithms for Bayesian quantile estimation of linear models, Tobit models, and regularized quantile regression.

In this chapter, we present a derivation of the GAL *pdf* from the mixture representation and both introduce and derive the cumulative distribution function (*cdf*) and the moment-generating function (*mgf*) of the GAL distribution. The GAL density and the GAL *cdf* are utilized to introduce an estimation method for the flexible Bayesian quantile regression in ordinal (FBQROR) models. Estimation of ordinal models, unlike linear models, is more challenging since identification restrictions and sampling of cut-points have to satisfy the ordering constraints. Moreover, through careful transformation of the mixture variables and joint sampling of the scale and shape parameters, we are able to achieve low autocorrelation in our MCMC draws. This result is a substantial improvement compared to the extremely high autocorrelation reported by Yan and Kottas (2017). Our sampling scheme can therefore improve the algorithm for Bayesian quantile regression in linear, Tobit, and regularized regression models as presented by Yan and Kottas (2017).

We illustrate the proposed methodology in two simulation studies where the errors are generated from a symmetric (logistic) distribution and an asymmetric (chi-square) distribution. The results show that the FBQROR model can maintain the actual skewness of the data across all considered quantiles. Furthermore, the FBQROR models can provide better model fit compared with the fit obtained from Bayesian quantile regression in ordinal (BQROR) models assuming an AL distribution. Finally, we implement our FBQROR model in an application related to the recent housing crisis and the Great Recession (December 2007–June 2009). Specifically, we analyze how various socioeconomic and demographic factors and exposure to financial distress are associated with differences in views on the financial benefits of homeownership following

the Great Recession. The results increase our understanding and offer new insights which may be important for policymakers and researchers interested in the US housing market.

The remainder of the chapter is organized as follows. Section 2 presents some fundamental properties of the GAL distribution. Section 3 presents the FBQROR model and its estimation procedure. Section 4 illustrates the algorithm in two simulation studies, and Section 5 implements the algorithm to examine US public opinion on homeownership. Section 6 presents some concluding remarks.

2. THE GAL DISTRIBUTION

The GAL distribution is obtained by introducing a shape parameter into the mean of the normal kernel in the normal-exponential mixture representation of the AL distribution and mixing with respect to a half-normal distribution. This hierarchical representation allows the skewness and mode to vary for a given quantile/percentile and hence provides the much-needed flexibility for Bayesian quantile regression.

Suppose Y is a random variable that has the following mixture representation,

$$Y = \mu + \sigma A W + \sigma \alpha S + \sigma [BW]^{\frac{1}{2}} U, \tag{1}$$

where $W \sim \mathcal{E}(1), S \sim N^{+}(0,1), U \sim N(0,1), A \equiv A(p) = \frac{1-2p}{p(1-p)}$ and $B \equiv B(p) = \frac{2}{p(1-p)}$. Here, $\mathcal{E}, N^{+}$, and N denote exponential, half-normal, and normal distributions, respectively. Then, Y follows a GAL distribution denoted $Y \sim GAL\ (\mu, \sigma, p, \alpha)$ and has the *pdf*

$$\begin{aligned} f(y|\theta) \ = & \frac{2p(1-p)}{\sigma} \left(\left[\Phi\left(\frac{y^*}{\alpha} - \alpha p_{\alpha_-} \right) - \Phi\left(-\alpha p_{\alpha_-} \right) \right] \right. \\ & \exp\left\{ -y^* p_{\alpha_-} + \frac{1}{2} (\alpha p_{\alpha_-})^2 \right\} \times I\left(\frac{y^*}{\alpha} > 0 \right) \\ & \left. + \Phi\left(\alpha p_{\alpha_+} - \frac{y^*}{\alpha} I\left(\frac{y^*}{\alpha} > 0 \right) \right) \exp\left\{ -y^* p_{\alpha_+} + \frac{1}{2} (\alpha p_{\alpha_+})^2 \right\} \right), \end{aligned} \tag{2}$$

where $\theta = (\mu, \sigma, p, \alpha)$, $y^* = (y - \mu)/\sigma$, μ is the location parameter, σ is the scale parameter, α is the shape parameter, $p_{\alpha_+} = p - I(\alpha > 0)$ and $p_{\alpha_-} = p - I(\alpha < 0)$ with $p \in (0,1)$. The derivation of the GAL *pdf* from the hierarchical representation is presented in Appendix A.1 and largely follows the notations used by Yan

and Kottas (2017). Note that when $\alpha = 0$, the GAL *pdf* reduces to the *pdf* of an AL distribution.

We explore the GAL distribution in greater detail and propose the *cdf* and *mgf* of the GAL distribution. The *cdf* denoted by F can be compactly written as,

$$\begin{aligned} F(y\,|\,\theta) = & \left(1 - 2\Phi\left(-\frac{y^*}{|\alpha|}\right) - \frac{2p(1-p)}{p_{\alpha_-}}\exp\left\{-y^* p_{\alpha_-} + \frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}\right. \\ & \left.\left[\Phi\left(\frac{y^*}{\alpha} - \alpha p_{\alpha_-}\right) - \Phi(-\alpha p_{\alpha_-})\right]\right) I\left(\frac{y^*}{\alpha} > 0\right) \\ & + I(\alpha < 0) - \frac{2p(1-p)}{p_{\alpha_+}}\exp\left\{-y^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\} \\ & \times \Phi\left[\alpha p_{\alpha_+} - \frac{y^*}{\alpha} I\left(\frac{y^*}{\alpha} > 0\right)\right], \end{aligned} \tag{3}$$

and the *mgf* denoted by $M_Y(t)$ has the following expression,

$$M_Y(t) = 2p(1-p)\left[\frac{(p_{\alpha_+} - p_{\alpha_-})}{(p_{\alpha_-} - \sigma t)(p_{\alpha_+} - \sigma t)}\right]\exp\left\{\mu t + \frac{1}{2}\alpha^2\sigma^2 t^2\right\}\ \Phi(|\alpha|\sigma t). \tag{4}$$

Both the *cdf* and *mgf* have been derived and presented in Appendix A.2 and Appendix A.3, respectively. In addition, Appendix A.3 utilizes the *mgf* (Eq. 4) to derive the mean, variance, and skewness of the distribution. These distributional characteristics are extremely important for better understanding of the GAL distribution and for further development of flexible Bayesian quantile regression.

However, the GAL density given by Eq. (2) has the limitation that the parameter p no longer corresponds to the cumulative probability at the quantile for $\alpha \neq 0$. We let $\gamma = [I(\alpha > 0) - p]\,|\alpha|$ and re-express the mixture representation shown in Eq. (1) as follows,

$$Y = \mu + \sigma A W + \sigma C\,|\gamma|\,S + \sigma[BW]^{\frac{1}{2}}U, \tag{5}$$

where $C = [I(\gamma > 0) - p]^{-1}$. This re-parametrization yields the quantile-fixed GAL distribution that has the following *pdf*:

$$f_{p_0}(y|\eta) = \frac{2p(1-p)}{\sigma}\left(\left[\Phi\left(-y^*\frac{p_{\gamma_+}}{|\gamma|}+\frac{p_{\gamma_-}}{p_{\gamma_+}}|\gamma|\right)-\Phi\left(\frac{p_{\gamma_-}}{p_{\gamma_+}}|\gamma|\right)\right]\right.$$
$$\exp\left\{-y^*p_{\gamma_-}+\frac{\gamma^2}{2}\left(\frac{p_{\gamma_-}}{p_{\gamma_+}}\right)^2\right\}$$
$$\left.\times\ I\left(\frac{y^*}{\gamma}>0\right)+\Phi\left(-|\gamma|+y^*\frac{p_{\gamma_+}}{|\gamma|}I\left(\frac{y^*}{\gamma}>0\right)\right)\exp\left\{-y^*p_{\gamma_+}+\frac{\gamma^2}{2}\right\}\right), \tag{6}$$

where $\eta = (\mu, \sigma, \gamma)$, $p \equiv p(\gamma, p_0) = I(\gamma<0) + [p_0 - I(\gamma<0)]/g(\gamma)$, $p_{\gamma_+} = p - I(\gamma>0)$ and $p_{\gamma_-} = p - I(\gamma<0)$. The function $g(\gamma) = 2\Phi(-|\gamma|)\exp(\gamma^2/2)$ and $\gamma \in (L, U)$, where L is the negative square root of $g(\gamma) = 1 - p_0$ and U is the positive square root of $g(\gamma) = p_0$. The term quantile-fixed suggests that integration of GAL *pdf* (Eq. 6) to the upper limit μ equals p_0, so for regression purpose we can fix the quantile. The *cdf* for the quantile-fixed GAL density (Eq. 6) can be analogously derived as in Appendix A.2 to yield the following expression,

$$F_{p_0}(y|\eta) = \left(1-2\Phi\left(y^*\frac{p_{\gamma_+}}{\gamma}\right)+2p_{\gamma_+}\exp\left\{-y^*p_{\gamma_-}+\frac{\gamma^2}{2}\left(\frac{p_{\gamma_-}}{p_{\gamma_+}}\right)^2\right\}\right.$$
$$\left.\left[\Phi\left(-y^*\frac{p_{\gamma_+}}{|\gamma|}+\frac{p_{\gamma_-}}{p_{\gamma_+}}|\gamma|\right)-\Phi\left(\frac{p_{\gamma_-}}{p_{\gamma_+}}|\gamma|\right)\right]\right)I\left(\frac{y^*}{\gamma}>0\right)$$
$$+\ I(\gamma<0)+2p_{\gamma_-}\exp\left\{-y^*p_{\gamma_+}+\frac{\gamma^2}{2}\right\}$$
$$\left.\times\ \Phi\left(-|\gamma|+y^*\frac{p_{\gamma_+}}{|\gamma|}I\left(\frac{y^*}{\gamma}>0\right)\right)\right).$$

The quantile-fixed *cdf* (Eq. 7) is required for constructing the likelihood of the FBQROR model and plays a critical role in the MCMC sampling of the scale parameter, shape parameter, and cut-points or thresholds.

To better discern the GAL distribution, Fig. 1 presents a graphical comparison between the quantile-fixed GAL and AL *pdf*'s for three different quantiles. We observe that the GAL distribution, unlike the AL distribution, allows the mode to vary rather than being fixed at $\mu = 0$ at all quantiles. Besides, the GAL distribution can be positively or negatively skewed at all quantiles depending on the value of γ. For example, at the median $p_0 = 0.50$ the GAL distribution is positively skewed for $\gamma<0$ and negatively skewed for $\gamma>0$. Also, the GAL

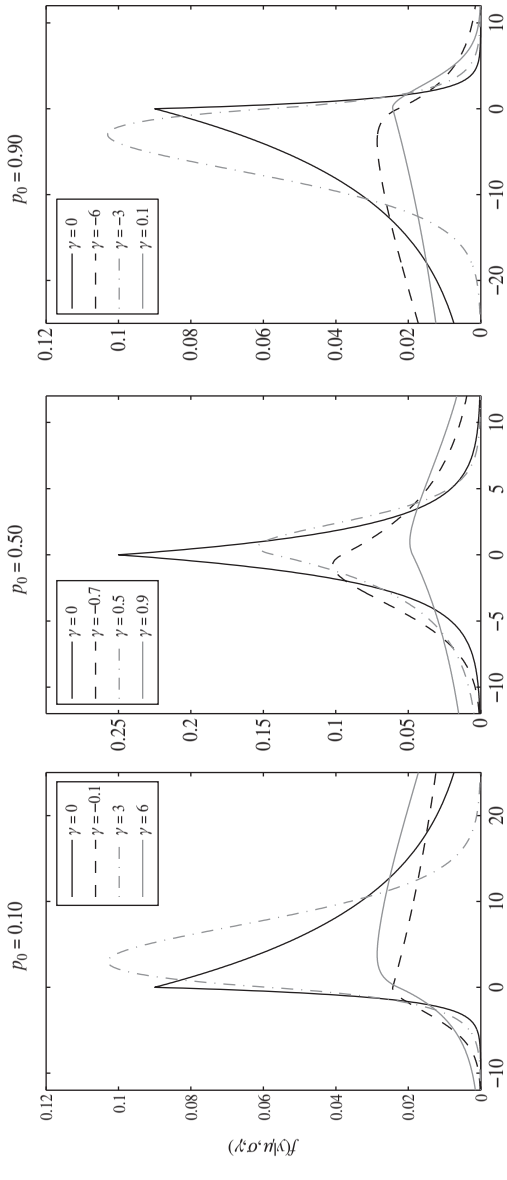

Fig. 1. Probability Density Plots of the AL ($\gamma = 0$) and the GAL ($\gamma \neq 0$) Distributions.

distribution can have tails that are heavier or narrower than the AL distribution. These characteristics make the GAL distribution more flexible than the AL distribution.

3. THE FBQROR MODEL

Ordinal models arise when the dependent (response) variable is discrete, and the outcomes are inherently ordered or ranked such that the scores assigned to outcomes have an ordinal meaning, but no cardinal interpretation (Jeliazkov & Rahman, 2012; Johnson & Albert, 2000). For example, in a survey on public opinion to allow more offshore drilling, responses may be recorded as follows: 1 for "strongly oppose," 2 for "somewhat oppose," 3 for "somewhat support," and 4 for "strongly support" (Mukherjee & Rahman, 2016). These responses have ordinal meaning but no cardinal interpretation. Therefore, one cannot say that a score of 4 implies four times more support compared with a score of 1.

We adopt the latent variable approach and represent the FBQROR model using a continuous latent random variable z_i expressed as a function of covariates and error as

$$z_i = x_i'\beta + \varepsilon_i, \quad \forall\, i = 1, \ldots, n, \tag{7}$$

where x_i is a $k \times 1$ vector of covariates, β is a $k \times 1$ vector of unknown parameters at the p_0th quantile, ε_i follows a GAL distribution, that is, $\varepsilon_i \sim GAL(0, \sigma, \gamma)$ and n denotes the number of observations. Note that we have suppressed the dependence of parameters on p_0 for notational simplicity. The variable z_i is unobserved and relates to the observed discrete response y_i, which has J categories or outcomes, via the cut-point vector ξ as follows:

$$\xi_{j-1} < z_i \leq \xi_j \Rightarrow y_i = j, \qquad \forall\ i = 1, \ldots, n;\ j = 1, \ldots, J, \tag{8}$$

where $\xi_0 = -\infty$ and $\xi_J = \infty$. In addition, ξ_1 is typically set to 0, which anchors the location of the distribution required for parameter identification (see Jeliazkov, Graves, & Kutzbach, 2008). Given the data vector $y = (y_1, \ldots, y_n)'$, the likelihood for the model expressed as a function of unknown parameters $(\beta, \sigma, \gamma, \xi)$ can be written as,

$$\begin{aligned} f(\beta, \sigma, \gamma, \xi; y) &= \prod_{i=1}^{n} \prod_{j=1}^{J} P(y_i = j \mid \beta, \sigma, \xi, \gamma)^{I(y_i = j)} \\ &= \prod_{i=1}^{n} \prod_{j=1}^{J} \left[F_{p_0}\left(\frac{\xi_j - x_i'\beta}{\sigma} \right) - F_{p_0}\left(\frac{\xi_{j-1} - x_i'\beta}{\sigma} \right) \right]^{I(y_i = j)} \end{aligned} \tag{9}$$

where, $F_{p_0}(\cdot) \equiv F(\cdot \mid 0, 1, \gamma)$ denotes the *cdf* of the GAL distribution and $I(y_i = j)$ is an indicator function, which equals 1 if $y_i = j$ and 0 otherwise.

Working directly with the GAL distribution is difficult, so we replace the error term with its mixture representation (Eq. 5) and rewrite the FBQROR model as follows:

$$z_i = x_i'\beta + \sigma A w_i + \sigma C\,|\,\gamma\,|\,s_i + \sigma\sqrt{B w_i}\,u_i, \quad \forall\ i = 1,\ldots,n. \tag{10}$$

The above formulation (Eq. 10) implies that the latent variable $z_i\,|\,\beta, w_i, s_i, \sigma, \gamma \sim N(x_i'\beta + \sigma C\,|\,\gamma\,|\,s_i + \sigma A w_i, \sigma^2 B w_i)$. However, the presence of the scale parameter σ in the conditional mean is not conducive to the construction of MCMC algorithm (Kozumi & Kobayashi, 2011). We reparameterize and write the model as,

$$z_i = x_i'\beta + A\nu_i + C\,|\,\gamma\,|\,h_i + \sqrt{\sigma B \nu_i}\,u_i, \quad \forall\ \ i = 1,\ldots,n, \tag{11}$$

where $h_i = \sigma s_i$ and $\nu_i = \sigma w_i$, which in turn implies that $h_i \sim N^+(0, \sigma^2)$ and $\nu_i \sim \mathcal{E}(\sigma)$ for $i = 1,\ldots,n$. Both reformulations are necessary for computational efficiency and low autocorrelation in MCMC draws. Note that the first reparameterization was not utilized by Yan and Kottas (2017), and hence, our approach provides a better alternative to estimating linear, Tobit, and regularized lasso quantile regression.

Ordinal models present two additional challenges: location and scale restrictions for identification of the parameters and ordering constraints in sampling of cut-point vector ξ (see Jeliazkov et al., 2008; Rahman, 2016). In the FBQROR model, both location and scale restrictions are enforced by fixing two cut-points since the variance of a GAL distribution is not fixed due to its dependence on α even if we set $\sigma = 1$, as shown in Theorem 4 in Appendix A. The ordering constraint is resolved using the following logarithmic transformation:

$$\delta_j = \ln(\xi_{j+2} - \xi_{j+1}), \qquad 1 \le j \le J-3. \tag{12}$$

The original cut-points are then obtained using Eq. (12) by one-to-one mapping between $\delta = (\delta_1,\ldots,\delta_{J-3})'$ and $\xi = (\xi_3,\ldots,\xi_{J-1})'$, where ξ_2 is fixed at some constant c and recall that $\xi_0 = -\infty$, $\xi_1 = 0$, and $\xi_J = \infty$.

We next employ the Bayes' theorem and derive the joint posterior density as proportional to the product of the likelihood and prior distributions. We employ standard prior distributions as follows,

$$\begin{aligned} &\beta \sim N(\beta_0, B_0), \quad \sigma \sim IG(n_0/2,\ d_0/2), \\ &\gamma \sim SB(L, U), \quad \delta \sim N(\delta_0, D_0), \end{aligned} \tag{13}$$

where N, IG, and SB denote normal, inverse-gamma, and scaled-Beta distributions, respectively. The lower and upper bounds of the scaled-Beta distribution are obtained as mentioned in Section 2. Combining the likelihood and the prior distributions, the augmented joint posterior density can be written as,

$$\begin{aligned}\pi(z,\beta,\nu,h,\sigma,\gamma,\delta\,|\,y) &\propto f(y\,|\,z,\beta,\nu,h,\sigma,\gamma,\delta)\pi(z\,|\,\beta,\nu,h,\sigma,\gamma,\delta)\pi(\nu\,|\,\sigma)\pi(h\,|\,\sigma)\pi(\beta)\\ &\quad\times\pi(\sigma)\pi(\gamma)\pi(\delta)\\ &\propto\left\{\prod_{i=1}^{n}f(y_i\,|\,z_i,\beta,\nu_i,h_i,\sigma,\gamma,\delta)\pi(\nu_i\,|\,\sigma)\pi(h_i\,|\,\sigma)\right\}\pi(z\,|\,\beta,\sigma,\nu,h,\gamma)\\ &\quad\times\pi(\beta)\pi(\sigma)\pi(\gamma)\pi(\delta)\\ &\propto\left\{\prod_{i=1}^{n}f(y_i\,|\,z_i,\delta)\pi(\nu_i\,|\,\sigma)\pi(h_i\,|\,\sigma)\right\}\pi(z\,|\,\beta,\sigma,\nu,h,\gamma)\\ &\quad\times\pi(\beta)\pi(\sigma)\pi(\gamma)\pi(\delta),\end{aligned}\tag{14}$$

where the last line in the likelihood, based on $GAL(0,\ \sigma,\ \gamma)$, uses the fact that given z and δ, the observed y is independent of the remaining parameters, because Eq. (8) determines y_i given (z,δ) with probability 1. The conditional density of latent data z is obtained from Eq. (11) and is given by $\pi(z\,|\,\beta,\sigma,\nu,h,\gamma)=\prod_{i=1}^{n}N(z_i\,|\,x_i'\beta+A\nu_i+C\,|\,\gamma\,|\,h_i,\sigma B\nu_i)$. Additionally, the prior distributions for $(\beta,\sigma,\gamma,\delta)$ are assumed to be independent in Eq. (14). Using the preceding explanations, the "complete data posterior" in Eq. (14) can be expressed as,

$$\begin{aligned}\pi(z,\beta,\nu,h,\sigma,\gamma,\delta\,|\,y)\ \propto&\left\{\prod_{i=1}^{n}1\{\xi_{y_i-1}<z_i\le\xi_{y_i}\}N(z_i\,|\,x_i'\beta+A\nu_i+C\,|\,\gamma\,|\,h_i,\sigma B\nu_i)\right.\\ &\left.\times\mathcal{E}(\nu_i\,|\,\sigma)\,N^{+}(h_i\,|\,0,\,\sigma^2)\right\}N(\beta\,|\,\beta_0,B_0)\,IG(\sigma\,|\,n_0/2,\,d_0/2)\\ &\times SB(\gamma\,|\,L,\,U)\,N(\delta\,|\,\delta_0,D_0).\end{aligned}\tag{15}$$

The objects of interest, that is, $(z,\beta,\nu,h,\sigma,\gamma,\delta)$ can be sampled by deriving the conditional posterior densities from the complete data posterior (Eq. 15) and judiciously using the full likelihood (Eq. 9) as presented in Algorithm 1. We note that our proposed algorithm is a form of MH within partially collapsed Gibbs sampler, and care has been taken to guarantee convergence to stationary distribution as given by van Dyk and Jiao (2015).

Algorithm 1. Sampling in FBQROR Model.

- Sample $\beta\,|\,z,\nu,h,\sigma,\gamma\sim N(\tilde{\beta},\tilde{B})$, where

$$\tilde{B}^{-1}=\left(B_0^{-1}+\sum_{i=1}^{n}\frac{x_ix_i'}{\sigma B\nu_i}\right)\quad\text{and}$$

$$\tilde{\beta}=\tilde{B}\left(\sum_{i=1}^{n}\frac{x_i(z_i-A\nu_i-C\,|\,\gamma\,|\,h_i)}{\sigma B\nu_i}+B_0^{-1}\beta_0\right).$$

- Sample (σ,γ) marginally of (z,ν,h) using a joint random-walk MH algorithm. The proposed values (σ',γ') are generated from a truncated bivariate

normal distribution $TBN_{(0,\infty)\times(L,U)}((\sigma_c,\gamma_c),\iota_1^2\hat{D}_1)$, where (σ_c,γ_c) denote the current values, ι_1 denotes the tuning factor, and $\hat{D}_1$ is the negative inverse of the Hessian obtained by maximizing the log-likelihood (Eq. 9) with respect to (σ,γ). The proposed draws are accepted with MH probability,

$$\alpha_{MH}(\sigma_c,\gamma_c;\sigma',\gamma') = \min\left\{0, \ln\left[\frac{f(y\,|\,\beta,\sigma',\gamma',\delta)\pi(\beta,\sigma',\gamma',\delta)}{f(y\,|\,\beta,\sigma_c,\gamma_c,\delta)\pi(\beta,\sigma_c,\gamma_c,\delta)}\frac{\pi(\sigma_c,\gamma_c\,|\,(\sigma',\gamma'),\iota_1^2\hat{D}_1)}{\pi(\sigma',\gamma'\,|\,(\sigma_c,\gamma_c),\iota_1^2\hat{D}_1)}\right]\right\},$$

else, repeat (σ_c,γ_c) in the next MCMC iteration. Here, $f(\cdot)$ represents the full likelihood (Eq. 9) obtained as the difference of *cdf*, $\pi(\beta,\sigma,\delta,\gamma)$ denotes the prior distributions (Eq. 13), and $\pi(\sigma_c,\gamma_c\,|\,(\sigma',\gamma'),\iota_1^2\hat{D}_1)$ stands for the truncated bivariate normal probability with mean (σ',γ') and covariance $\iota_1^2\hat{D}_1$. The term $\pi(\sigma',\gamma'\,|\,(\sigma_c,\gamma_c),\iota_1^2\hat{D}_1)$ has an analogous interpretation.

- Sample $\nu_i\,|\,z_i,\beta,h,\sigma,\gamma \sim GIG(0.5,\ a_i,\ b)$, for $i=1,\ldots,n$, where

$$a_i = \frac{(z_i - x_i'\beta - C\,|\,\gamma\,|\,h_i)^2}{\sigma B} \quad \text{and} \quad b = \left(\frac{A^2}{\sigma B} + \frac{2}{\sigma}\right).$$

- Sample $h_i\,|\,z_i,\beta,\nu_i,\sigma,\gamma \sim N^+(\mu_{h_i},\sigma_{h_i}^2)$ for $i=1,\ldots,n$, where

$$(\sigma_{h_i}^2)^{-1} = \left(\frac{1}{\sigma^2} + \frac{C^2\gamma^2}{\sigma B\nu_i}\right) \quad \text{and} \quad \mu_{h_i} = \sigma_{h_i}^2\left(\frac{C\,|\,\gamma\,|\,(z_i - x_i'\beta - A\nu_i)}{\sigma B\nu_i}\right).$$

- Sample $\delta\,|\,\beta,\sigma,\gamma,y$ marginally of (z,ν,h) using a random-walk MH step. The proposed value δ' is generated as $\delta' = \delta_c + u$, where $u \sim N(0_{J-3},\iota_2^2\hat{D}_2)$, ι_2 is a tuning parameter, and $\hat{D}_2$ is analogous to $\hat{D}_1$. Accept δ' with MH probability,

$$\alpha_{MH}(\delta_c,\delta') = \min\left\{0, \ln\left[\frac{f(y\,|\,\beta,\sigma,\gamma,\delta')\pi(\beta,\sigma,\gamma,\delta')}{f(y\,|\,\beta,\sigma,\gamma,\delta_c)\pi(\beta,\sigma,\gamma,\delta_c)}\right]\right\},$$

else, repeat δ_c. Again $f(\cdot)$ denotes the full likelihood (Eq. 9) and $\pi(\beta,\sigma,\delta,\gamma)$ denotes the priors.

- Sample $z_i\,|\,y,\beta,\nu_i,h_i,\sigma,\gamma,\delta \sim TN_{(\xi_{j-1},\xi_j)}(x_i'\beta + A\nu_i + C\,|\,\gamma\,|\,h_i,\sigma B\nu_i)$ for $i=1,2,\ldots,n$, where ξ is obtained from δ by one-to-one mapping using Eq. (12).

Starting with the regression coefficients, β is sampled from a multivariate normal distribution, draws from which are programmed in most statistical software. The scale and shape parameters (σ,γ) are jointly sampled, marginally of (z,ν,h), using a random-walk Metropolis–Hastings (MH) algorithm with proposals drawn from a truncated bivariate normal distribution. Joint sampling (together

with the transformations $h_i = \sigma s_i$ and $\nu_i = \sigma\omega_i$) is crucial for reducing the high autocorrelation in MCMC draws as observed by Yan and Kottas (2017). The latent weight ν follows a generalized inverse-Gaussian (GIG) distribution, draws from which are obtained using the technique proposed by Devroye (2014). Alternatively, one may employ the ratio of uniforms method or the envelope rejection method (Dagpunar, 1988, 1989, 2007). The mixture variable h is sampled from a half-normal distribution. Typical to ordinal models, the cut-points δ do not have a tractable distribution and is sampled marginally of (z, ν, h) using a random-walk MH algorithm (see Jeliazkov et al., 2008; Rahman, 2016). Finally, the latent variable z, conditional on the remaining parameters, is sampled from a truncated normal distribution (Botev, 2017). The derivations of the conditional posteriors and details of the MH algorithms are presented in Appendix B.

4. SIMULATION STUDIES

This section presents two simulation studies to demonstrate the performance of the proposed algorithm and illustrate the advantages of the FBQROR model compared to the BQROR model.

4.1. Simulation Study 1

In this simulation study, we estimate and compare the FBQROR model with the BQROR model when errors are generated from a symmetric distribution. Specifically, 300 observations are generated from the model $z_i = x_i'\beta + \varepsilon_i$, where covariates are sampled from a standard uniform distribution $Unif[0, 1]$, $\beta = (2, -3, 4)'$, and ε are sampled from a logistic distribution $\mathcal{L}(0, \pi^2/3)$. The resulting continuous variable z is symmetric and is utilized to construct the discrete response variable y based on the cut-point vector $\xi = (0, 2, 4)$. In our simulated data, the number of observations corresponding to the four categories of y is 42 (14%), 81 (27%), 99 (33%), and 78 (26%), respectively.

The posterior estimates of the parameters for FBQROR model are obtained based on the simulated data and the following moderately diffuse priors: $\beta \sim N(0_3, 10I_3)$, $\sigma \sim IG(5/2, 8/2)$, $\gamma \sim SB(L, U, 4, 4)$, and $\delta \sim N(0_{J-3}, I_{J-3})$ for $p_0 = (0.25, 0.5, 0.75)$, where (L, U) depends on the value of p_0 as mentioned in Section 2. Table 1 reports the MCMC results obtained from 15,000 iterations, after a burn-in of 5,000 iterations, along with the inefficiency factors calculated using the batch-means method (Greenberg, 2012). The parameters (σ, γ) are jointly sampled using random-walk MH algorithm with tuning parameters $\iota_1 = (\sqrt{1.7}, \sqrt{2.25}, \sqrt{2.0})$ to achieve an acceptance rate of approximately 33 percent for the three considered quantiles. Similarly, δ is sampled using a random-walk MH algorithm with tuning factor $\iota_2 = (\sqrt{4.0}, \sqrt{3.2}, \sqrt{2.5})$ to obtain an acceptance rate of around 33 percent. Inefficiency factors for all the model parameters are low which imply low autocorrelation in MCMC draws and trace plots of the MCMC iterations, as exhibited in Fig. 2 for the 25th quantile, display quick convergence. Trace plots for the other two quantiles are similar and

Table 1. Posterior Mean (MEAN), Standard Deviation (STD) and Inefficiency Factor (IF) of the Parameters in Simulation Study 1.

	FBQROR Model								
	25th Quantile			50th Quantile			75th Quantile		
$(\beta, \sigma, \gamma, \delta)$	Mean	STD	IF	Mean	STD	IF	Mean	STD	IF
β_1	1.07	0.32	2.94	2.17	0.31	2.66	3.23	0.35	4.91
β_2	−3.22	0.50	4.02	−3.14	0.48	3.67	−3.13	0.47	4.72
β_3	3.93	0.53	4.08	3.86	0.51	3.82	3.90	0.50	5.28
σ	0.64	0.10	3.38	0.75	0.09	4.30	0.60	0.09	4.10
γ	1.14	0.27	2.55	−0.06	0.17	3.76	−1.18	0.24	2.93
δ_1	0.73	0.14	4.75	0.71	0.15	4.84	0.67	0.14	5.87
	BQROR Model								
	25th Quantile			50th Quantile			75th Quantile		
(β, σ, δ)	Mean	STD	IF	Mean	STD	IF	Mean	STD	IF
β_1	1.15	0.30	2.46	2.12	0.29	2.56	2.86	0.26	4.43
β_2	−3.23	0.46	3.21	−2.96	0.43	3.32	−2.38	0.46	5.79
β_3	3.82	0.44	3.11	3.64	0.47	3.71	3.01	0.49	6.46
σ	0.65	0.07	3.55	0.71	0.07	3.84	0.45	0.05	5.92
δ_1	0.85	0.13	3.84	0.60	0.14	4.51	0.35	0.15	6.55

Note: The first panel presents results from the FBQROR model and the second panel presents results from the BQROR model.

have not been shown for the sake of brevity. The sampler is reasonably quick and takes approximately 160 seconds per 1,000 iterations.

The first panel of Table 1 presents the results for the FBQROR model. The results show that the posterior means for β are close to the true parameter values, the posterior mean of σ adjusts the scale of the distribution, and the posterior mean of δ_1 yields a value of ξ_3 close to 4, the true value used to generate the data. We also estimate the BQROR model using a modification of Algorithm 1 presented in Rahman (2016) − by fixing the second cut-point and introducing a scale parameter in the model. Prior distributions for (β, σ, δ) are identical to those of the FBQROR model. The results, presented in the second panel of Table 1, show that the posterior estimates for (β, σ, δ) are similar. Moving to the skewness parameter in the FBQROR model, the posterior mean of γ at $p_0 = (0.25, 0.50, 0.75)$ are $(1.14, -0.06, -1.18)$, which corresponds to a skewness of $(0.01, 0.20, 0.04)$, respectively. Note that the posterior mean of γ at $p_0 = 0.50$ is statistically equivalent to zero. These skewness values imply that the (latent) response variable is approximately symmetric at all considered quantiles, which is reassuring since our data were generated from a symmetric distribution. In contrast, the corresponding skewness values for the BQROR model are

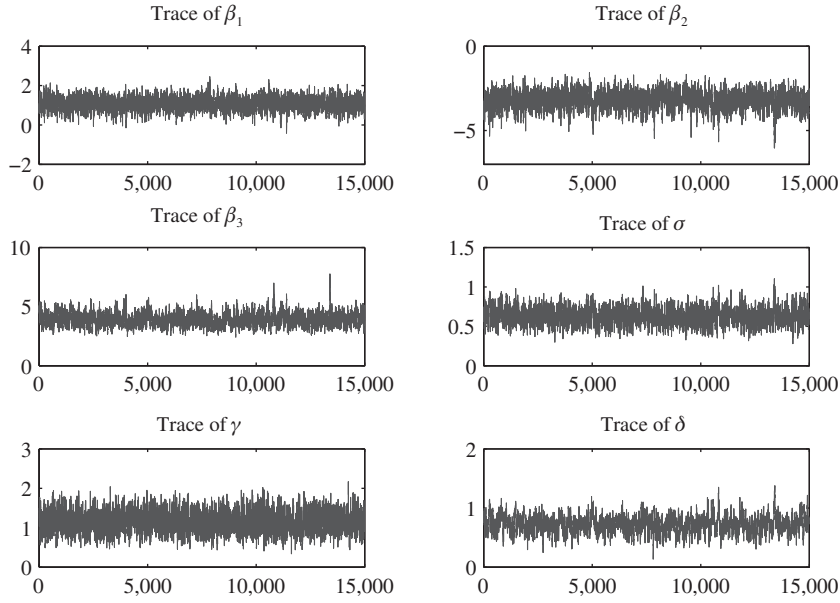

Fig. 2. Trace Plots of the MCMC Draws at the 25th Quantile for Simulation Study 1.

Table 2. Model Comparison using the Conditional Log-likelihood (ln L), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) in Simulation Study 1.

	25th Quantile (ln L, AIC, BIC)	50th Quantile (ln L, AIC, BIC)	75th Quantile (ln L, AIC, BIC)
FBQROR	(−338,688,710)	(−340,691,714)	(−338,688,710)
BQROR	(−346,701,720)	(−340,690,708)	(−348,706,724)

(1.64, 0, –1.64). Hence, the BQROR model fails to accommodate the symmetric characteristic of the data at the 25th and 75th quantiles.

We next investigate model fitness at different quantiles since various choices of quantile p_0 may be interpreted as corresponding to a different link function. Table 2 presents the conditional log-likelihood, the Akaike information criterion (AIC; Akaike, 1974), and the Bayesian information criterion (BIC, Schwarz, 1978) for both the FBQROR and BQROR models. Higher conditional log-likelihood is preferable, while lower values of AIC/BIC indicate a better model fit. As illustrated in Table 2, the conditional log-likelihood for the FBQROR model is identical to that for the BQROR model at the median, but higher at the other two considered quantiles. However, the FBQROR model has an extra shape parameter, and so to rule out the possibility of higher log-likelihood

arising due to additional parameters (i.e., overfitting), we compare the models using AIC and BIC. These two measures introduce different penalty terms to account for the number of model parameters. Based on AIC/BIC, there is strong evidence that the FBQROR model provides a better fit at the 25th and 75th quantiles, but there is some evidence in favor of BQROR model at the 50th quantile. The poor fit of the BQROR model at the first and third quartiles reflects the rigidity of the AL distribution, since $p_0 = 0.25$ (0.75) forces the AL distribution to be positively (negatively) skewed.

4.2. Simulation Study 2

Once again we estimate the FBQROR and BQROR models with simulated data, but now the errors are generated from a chi-square distribution such that the resulting distribution for the continuous latent variable z is positively skewed. In particular, 300 observations are generated from the model $z_i = x_i'\beta + \varepsilon_i$, where covariates are sampled from a standard uniform distribution $Unif[0, 1]$, $\beta = (3, -7, 5)'$ and ε is generated from $\chi^2(4) - 4$, that is, a demeaned chi-square distribution. The discrete response variable y is obtained from z based on cut-point vector $\xi = (0, 3, 6)$, which yields 74 (24.67%), 110 (36.67%), 65 (21.67%), and 51 (17.00%) observations in the four categories of y.

Table 3 reports the MCMC estimates obtained from 15,000 iterations after a burn-in of 5,000 iterations. Identical prior distributions as in the first simulation study were used for both FBQROR and BQROR models. The parameters (σ, γ) and δ are sampled using random-walk MH algorithm with tuning factors $\iota_1 = (\sqrt{0.3}, \sqrt{0.7}, \sqrt{1.45})$ and $\iota_2 = (\sqrt{3.25}, \sqrt{3.1}, \sqrt{2.75})$ to achieve an acceptance rate of approximately 33 percent. The inefficiency factors are low and trace plots, as displayed in Fig. 3 for the 50th quantile, show quick convergence. Trace plots at the other two quantiles are similar. Computational time remains unchanged at approximately 160 seconds per 1,000 iterations.

Table 3 presents the results for the FBQROR and BQROR models in the first and second panels, respectively. In the FBQROR model, the posterior estimates of β are close to the true values $(3, -7, 5)$ and the posterior estimates of σ adjusts to capture the spread. Moreover, the posterior estimates of γ capture the skewness of the data extremely well. Specifically, the posterior mean of γ is not statistically different from zero at the 25th quantile, so skewness is approximately same as that of BQROR model (1.64). However, the skewness at the 50th (75th) quantile is 1.31 (0.16) compared to a skewness of 0 (-1.64) in the BQROR model. Hence, the FBQROR model correctly captures the positive skewness of the simulated data at all quantiles. Model comparison also points to the superiority of the FBQROR model as seen in Table 4. The conditional log-likelihood for the FBQROR model is higher than that of BQROR model at the 50th and 75th quantiles, but identical at the 25th quantile. The AIC and BIC values suggest that there is strong evidence to select the FBQROR model at the 50th and 75th quantiles, and some evidence to favor the BQROR model at the 25th quantile. Once again, the flexibility offered by the FBQROR model in

Table 3. Posterior Mean (MEAN), Standard Deviation (STD), and Inefficiency Factor (IF) of the Parameters in Simulation Study 2.

	FBQROR Model								
	25th Quantile			50th Quantile			75th Quantile		
$(\beta, \sigma, \gamma, \delta)$	Mean	STD	IF	Mean	STD	IF	Mean	STD	IF
β_1	1.60	0.36	2.87	2.83	0.37	3.21	4.50	0.44	4.00
β_2	−6.50	0.57	3.37	−6.39	0.61	3.98	−6.15	0.64	3.47
β_3	3.69	0.52	3.01	3.59	0.54	3.36	3.44	0.58	3.24
σ	0.74	0.11	8.08	0.74	0.12	2.25	0.79	0.10	2.66
γ	0.09	0.15	4.75	−0.49	0.09	4.37	−1.33	0.17	2.47
δ_1	0.91	0.14	2.20	0.86	0.14	2.89	0.73	0.13	3.21
	BQROR Model								
	25th Quantile			50th Quantile			75th Quantile		
(β, σ, δ)	Mean	STD	IF	Mean	STD	IF	Mean	STD	IF
β_1	1.10	0.23	2.72	1.91	0.24	2.44	2.68	0.24	2.78
β_2	−4.40	0.40	3.71	−3.88	0.38	2.84	−3.26	0.41	3.20
β_3	2.49	0.33	3.01	2.11	0.35	2.75	1.80	0.34	2.83
σ	0.47	0.04	3.07	0.63	0.06	2.93	0.49	0.04	2.94
δ_1	0.51	0.15	3.23	0.25	0.14	3.02	−0.04	0.14	3.45

Note: The first panel presents results from the FBQROR model and the second panel presents results from the BQROR model.

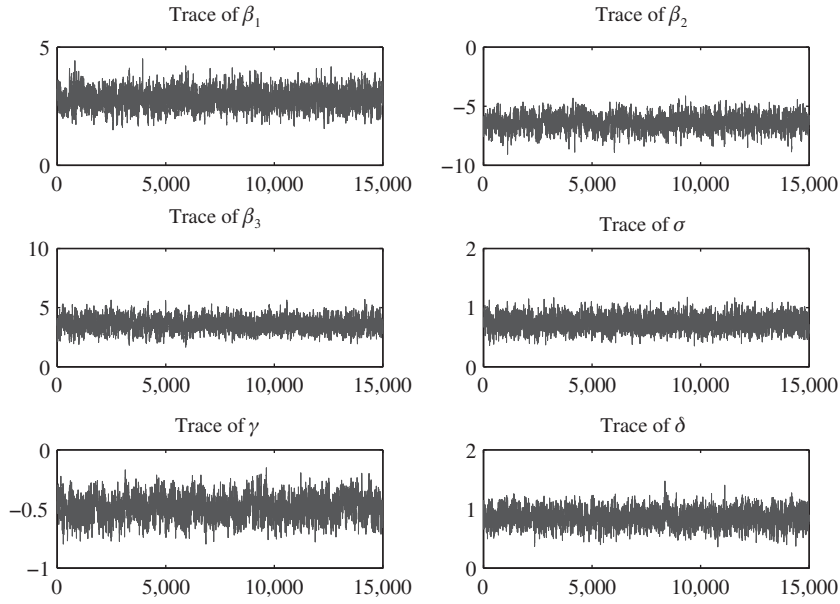

Fig. 3. Trace Plots of the MCMC Draws at the 50th Quantile for Simulation Study 2.

Table 4. Model Comparison using the Conditional Log-likelihood (ln L), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) in Simulation Study 2.

	25th Quantile (ln L, AIC, BIC)	50th Quantile (ln L, AIC, BIC)	75th Quantile (ln L, AIC, BIC)
FBQROR	(−318,649,671)	(−321,654,676)	(−333,678,700)
BQROR	(−318,646,664)	(−331,673,692)	(−358,727,745)

terms of modeling the skewness helps provide a better model fit compared to the rigid BQROR model.

5. APPLICATION

In the United States, consumers have typically viewed homeownership as a good long-term financial investment. However, the recent housing crisis and the subsequent economic recession (December 2007−June 2009) had substantial adverse effects, particularly on homeowners. Housing values depreciated considerably, more than four million foreclosures took place between 2008−2011, and one-fourth of all homeowners were living in houses worth less than the mortgage at the peak of the crisis (Belsky, 2013). Homeownership rate declined from 69.2 percent during Q2, 2004 to 66.4 percent during Q1, 2011 and further to 62.9 percent during Q2, 2016 (US Bureau of the Census). These adverse effects may have fundamentally altered the perceived benefits of homeownership as a good long-term investment (see Rohe and Lindblad, 2014 for a conceptual model). Consequently, it is of considerable interest to analyze public opinion on homeownership as an investment and examine how socioeconomic factors, demographic variables, and exposure to financial distress affect public responses.

This chapter utilizes the Higher Education/Housing Survey data of March 2011, conducted by the Princeton Survey Research Associates International and sponsored by the Pew Social and Demographic Trends project. Interviews were conducted over the telephone between March 15 and 29, 2011 on a nationally representative sample of 2,142 adults living in the continental United States. After removing missing responses, we are left with a sample of 1,799 observations for our analysis. Our dependent variable is response to the statement, "Some people say that buying home is the best long-term investment in the United States. Do you strongly agree, somewhat agree, somewhat disagree or strongly disagree?" Responses are recorded into one of the four categories; however, we append the responses "strongly disagree" and "somewhat disagree" as the former category had less than five percent observations. The survey also collected information on a wide range of socioeconomic, demographic and geographic variables, some of which are used as covariates in the model. Table 5 presents the description and summary statistics of all covariates and the response variable utilized in the study.

Table 5. Descriptive Summary of the Variables.

Variable	Description	Mean	STD
Log age	Logarithm of age (in years)	3.71	0.44
Log income	Logarithm of the mid-point of income category (in US dollars)	10.68	0.95
Household size	Number of members in the household	2.92	1.66
		Count	Percent
Female	Indicator variable for female gender	925	51.42
Post-bachelor's	Respondent's highest qualification is masters, professional or doctorate	257	14.29
Bachelor's	Respondent's highest qualification is bachelor's	395	21.96
Below bachelor's	Respondent holds a two-year associate degree, went to some college with no degree, or attended technical, trade or vocational school after high school	551	30.63
HS and Below	Respondent is a high school (HS) graduate or below	596	33.92
Full-time	Works full time	849	47.19
Part-time	Works part time	266	14.79
Unemployed	Either unemployed, student or retired	684	38.02
White	Respondent is a white-American	1293	71.87
African-American	Respondent is an African-American	272	15.12
All other races	Respondent is an Asian, Asian-American or belongs to some other race	234	13.01
Northeast	Lives in the northeast region of United States	249	13.84
West	Lives in the west region of United States	408	22.68
South	Lives in the south region of United States	822	45.69
Midwest	Lives in the midwest region of United States	320	17.79
Fin-better	Financially better-off post the Great Recession	537	29.85
Fin-same	Financially equivalent pre and post the Great Recession	445	24.74
Fin-worse	Financially worse-off post the Great Recession	817	45.41
	Strongly disagree or somewhat disagree that homeownership is the best long-term investment (LTI) in United States	310	17.23
Opinion	Somewhat agree that homeownership is the best LTI	828	46.03
	Strongly agree that homeownership is the best LTI	661	36.74

The average age of the sampled individuals is 44.84 years with a standard deviation of 18.59 years. Information on family income in the survey is recorded as one of nine income categories: $<10k$, $10k-20k$, $20k-30k$, $30k-40k$, $40k-50k$, $50k-75k$, $75k-100k$, $100k-150k$, and $>150k$, where k denotes a thousand dollars and US $5,000 and US $1,70,000 have been imputed for the first- and last-income categories. We include the logarithm of the mid-point of the income category as a variable in the model. Mean household size is 2.92

with a standard deviation of 1.66. The percent of female is slightly more than that of males. Educational classification shows that HS and below forms the largest category (33.92%) and post-bachelors forms the smallest category (14.29%), with proportions decreasing as we move from the lowest to the highest educational category. Employment status shows that 61.98% are either employed full-time or part-time, while the remaining are unemployed, student, or retired. With respect to race, the sample is predominantly white (71.87%), followed by African-Americans (15.12%) and all other races (13.01%). Geographically, most of the sampled individuals live in the South (45.69%), followed by West (22.68%), Midwest (17.79%), and Northeast (13.84%). These regional classifications are as defined by the US Census Bureau. To measure exposure to financial distress, we include self-reported financial condition pre- and post-Great Recession. As expected, almost half the sampled individuals (45.41%) are financially worse-off post the Great Recession.

Moving to the response variable, Table 5 shows that more than three-fourths of the sampled individuals (82.77%) somewhat or strongly agree that homeownership is the best long-term investment. Therefore, US public opinion on homeownership remains largely unchanged even after the housing meltdown and the Great Recession. A similar conclusion has been obtained using data from the Survey of Consumers collected by the University of Michigan and the National Housing Survey collected by Fannie Mae (Belsky, 2013). This result is primarily due to the financial benefits of homeownership making owning more lucrative than renting, especially in the long run. Two related articles that have studied the preference for homeownership versus renting using binary models on survey data are those by Bracha and Jamison (2012) and Drew and Herbert (2013). Both the studies find no fundamental shifts in attitude toward homeownership.

We employ the FBQROR and BQROR models to analyze public opinion on homeownership as the best long-term investment based on the covariates presented in Table 5. The MCMC results, presented in Table 6, are based on 15,000 iterations after a burn-in of 5,000 iterations with identical priors as in the simulation studies. With three values of the ordinal response variable, we have two cut-points and they are fixed at $(0, 3)$ for both models across quantiles. Similar to the simulation studies, (σ, γ) is sampled using joint random-walk MH algorithm with tuning parameters $\iota_1 = (\sqrt{3.0}, \sqrt{2.4}, \sqrt{4.4})$ to get an acceptance rate of approximately 33 percent for the three considered quantiles. The inefficiency factors of the parameters are all less than five, and trace plots of MCMC draws, as displayed in Fig. 4 for the 75th quantile, show good mixing. Trace plots at the other two quantiles are similar.

The results presented in Table 6 clearly show that the posterior estimates from FBQROR and BQROR models are fairly similar across all quantiles. Hence, we restrict our attention to the FBQROR model and use the BQROR model for model comparison. Moreover, we primarily discuss the covariates which are statistically different from zero at the 95% probability level. As seen from Table 6, age has a positive effect which implies that older individuals are

Table 6. Posterior Mean (Mean) and Standard Deviation (STD) of the Parameters in the FBQROR and BQROR Models for the Homeownership Application.

	FBQROR						BQROR					
	25th Quantile		50th Quantile		75th Quantile		25th Quantile		50th Quantile		75th Quantile	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Intercept	−3.11	0.93	−1.72	1.02	0.18	0.90	−3.08	0.89	−1.43	0.93	1.44	0.62
Log age	0.52	0.17	0.60	0.17	0.55	0.16	0.18	0.16	0.63	0.17	0.47	0.12
Log income	0.18	0.08	0.16	0.09	0.15	0.08	0.25	0.08	0.14	0.08	0.04	0.05
Household size	0.01	0.04	0.01	0.04	0.01	0.04	−0.01	0.04	0.01	0.04	0.02	0.03
Female	0.64	0.14	0.62	0.14	0.54	0.13	0.57	0.13	0.57	0.13	0.32	0.09
Post-bachelor's	−0.83	0.22	−0.86	0.22	−0.81	0.20	−0.50	0.21	−0.86	0.21	−0.62	0.15
Bachelor's	−0.72	0.20	−0.74	0.19	−0.69	0.18	−0.43	0.18	−0.74	0.18	−0.52	0.13
Below bachelor's	−0.37	0.17	−0.36	0.17	−0.33	0.16	−0.26	0.16	−0.34	0.16	−0.25	0.12
Full-time	−0.04	0.16	−0.02	0.16	−0.02	0.14	−0.09	0.15	−0.01	0.15	0.01	0.10
Part-time	−0.07	0.21	−0.05	0.21	−0.09	0.19	−0.01	0.19	−0.06	0.20	−0.12	0.14
White	0.04	0.21	0.01	0.21	0.01	0.18	0.18	0.18	−0.03	0.20	−0.06	0.14
African-American	0.05	0.26	0.03	0.27	0.05	0.24	0.17	0.23	−0.01	0.25	−0.01	0.17
Northeast	0.37	0.24	0.34	0.24	0.27	0.22	0.41	0.22	0.28	0.23	0.12	0.16
West	0.44	0.22	0.46	0.22	0.34	0.19	0.39	0.19	0.42	0.21	0.17	0.14
South	0.34	0.19	0.35	0.19	0.23	0.17	0.38	0.17	0.30	0.18	0.05	0.12
Fin-worse	−0.45	0.16	−0.48	0.16	−0.43	0.14	−0.34	0.15	−0.46	0.15	−0.31	0.11
Fin-same	−0.25	0.18	−0.26	0.18	−0.24	0.17	−0.25	0.18	−0.25	0.17	−0.15	0.13
σ	0.89	0.13	1.07	0.05	0.83	0.08	0.88	0.03	1.06	0.04	0.59	0.02
γ	1.05	0.31	−0.10	0.10	−1.09	0.24	..	..	..	..	..	..

more likely to strongly agree that homeownership is the best long-term investment. This result is consistent with the view that older adults are less likely to change their attitude when faced with harsh economic experiences such as an economic crisis (Giuliano & Spilimbergo, 2014; Malmendier & Nagel, 2011). Our result also finds support in the study by Bracha and Jamison (2012), who find that older individuals are more confident about homeownership (relative to renting) following large price declines. Income has a positive effect implying that higher income individuals are more likely to agree with the investment benefits of homeownership. However, income is an important factor only at the 25th quantile. This result is somewhat consistent with that of Drew and Herbert

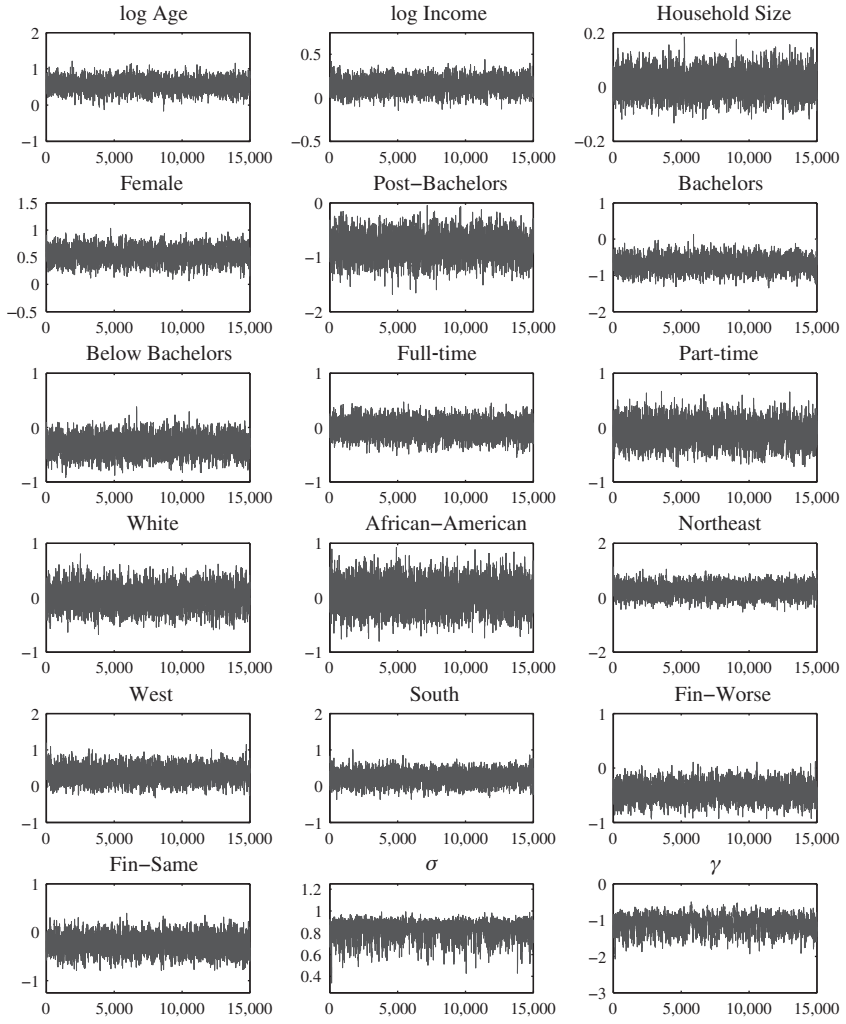

Fig. 4. Trace Plots of the MCMC Draws at the 75th Quantile for the Homeownership Application.

(2013), who find no statistically significant association between income and viewing homeownership as a better financial choice over renting.

Opinions across gender often vary due to risk perceptions, and this is reflected in our results. We find that females are more likely to strongly agree that homeownership is the best long-term investment across all quantiles. This result is consistent with the view that females are more risk averse than males, and homeownership has historically been a safe investment. However, our results are in contrast to those of Bracha and Jamison (2012), who find that

females are more uncertain about the financial gain from buying a house. Higher education has a negative effect on positive opinion about homeownership. The negative post-bachelors coefficient indicates that an individual with a post-bachelor's degree (relative to HS and below education) will be less willing to strongly agree that homeownership is the best long-term investment. Similarly, individuals with a bachelor's or below bachelor's education are less likely to positively view the investment benefits of homeownership. The negative effect of higher education on homeownership is also reported by Bracha and Jamison (2012) and Drew and Herbert (2013).

Employment status, whether full-time or part-time as compared to being unemployed, does not account for differences in opinion on the financial benefits of homeownership. The same is true for the white and African-American race indicators. Thus, individuals have similar views on homeownership as an investment irrespective of employment status or race. During the housing crisis, the decline in housing prices varied tremendously across geographic regions. West and South regions experienced the largest declines in housing prices. Accordingly, we include indicator variables for geographic regions to capture differences in opinion due to residing in different regions. The results from the lower quantiles suggest that individuals living in the West, relative to Midwest, are more likely to strongly agree on the financial benefits of homeownership. This result is interesting since people living in the West were the hardest hit in terms of housing price declines. Undoubtedly, the housing meltdown and the economic crisis caused serious financial distress to a large number of individuals in the United States. This experience may have altered views on homeownership. To capture the effect of financial distress on homeownership views, we include indicator variables for post-recession financial situation. The results indicate that individuals who are financially worse-off post the Great Recession, relative to those who are better-off, are less probable to strongly agree that homeownership is the best long-term investment. Hence, our results provide evidence that financial hardship endured during the Great Recession negatively impacted public views on the investment benefits of homeownership.

In the previous paragraphs, we discussed the direction of covariate effects on the last outcome, that is, strongly agree that homeownership is the best long-term investment. The direction of covariate effect on the first outcome (strongly disagree or somewhat disagree) is the opposite, while the effect on the second outcome (somewhat agree) cannot be known *a-priori*. This is because the link function in ordinal models is nonlinear, and hence, the regression coefficients do not give the covariate effects. To make it clear, we calculate the marginal effect for three variables: female, post-bachelor's, and worse financial condition (Jeliazkov & Vossmeyer, 2018). The change in predicted probabilities for the three response are reported in Table 7. We see that at the 25th quantile, individuals who are exposed to financial distress (i.e., financially worse-off) are 6.71% less likely to "strongly agree," 2.19% more likely to "somewhat agree," and 4.52% more likely to "strongly disagree or somewhat disagree" that homeownership is the best long-term investment. The marginal effect of financial distress on the responses "strongly agree" is more pronounced at the 50th and 75th

Table 7. Change in Predicted Probability of the Responses.

	Female			Post-bachelors			Fin-worse		
	25th Quantile	50th Quantile	75th Quantile	25th Quantile	50th Quantile	75th Quantile	25th Quantile	50th Quantile	75th Quantile
$\Delta P(y = 1)$	−0.0630	−0.0553	−0.0504	0.0962	0.0900	0.0852	0.0452	0.0429	0.0404
$\Delta P(y = 2)$	−0.0322	−0.0405	−0.0467	0.0174	0.0282	0.0473	0.0219	0.0302	0.0360
$\Delta P(y = 3)$	0.0952	0.0958	0.0970	−0.1136	−0.1182	−0.1325	−0.0671	−0.0731	−0.0764

Note: Somewhat or strongly disagree ($y = 1$), somewhat agree ($y = 2$), and strongly agree ($y = 3$) that homeownership is the best long-term investment.

Table 8. Model Comparison using the Conditional Log-likelihood (ln L), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) in the Homeownership Application.

	25th Quantile (ln L, AIC, BIC)	50th Quantile (ln L, AIC, BIC)	75th Quantile (ln L, AIC, BIC)
FBQROR	(−1816,3671,3775)	(−1815,3668,3772)	(−1816,3671,3775)
BQROR	(−1824,3684,3783)	(−1815,3665,3764)	(−1818,3673,3772)

quantiles. We can similarly interpret the change in predicted probabilities for female and post-bachelor's education on the three responses for different quantiles.

To assess model fitness across quantiles, we report the conditional log-likelihood, AIC, and BIC in Table 8. The log-likelihood for the FBQROR model across all quantiles is higher or similar to that obtained from the BQROR model. However, according to AIC, there is strong (weak) evidence in favor of the FBQROR model at the 25th (75th) quantile, but weak evidence in favor of the BQROR model at the 50th quantile. Based on BIC, there is strong evidence to prefer the FBQROR (BQROR) model at the 25th (50th) quantile, but positive evidence to prefer the BQROR model at the 75th quantile.

6. CONCLUSION

This chapter presents an estimation algorithm for Bayesian quantile regression in univariate ordinal models where the error is assumed to follow a GAL distribution, referred to as the FBQROR model. To propose this estimation procedure, we explore the GAL distribution and introduce and derive its cumulative distribution function and moment generating function. We show that the advantages offered by the GAL distribution – which allows the mode, skewness and tails to vary for any given quantile – can be gainfully utilized to better estimate Bayesian quantile regression in ordinal models. We also emphasize on the efficiency of the MCMC algorithm, which is attained through suitable transformation of the variables and joint sampling of the scale and shape parameters. The practical advantages of the proposed model are illustrated in multiple simulation

studies via model comparison, where it is observed that the FBQROR model can provide a better model fit compared to an ordinal model with an AL distribution, labeled BQROR model (Rahman, 2016). Our proposed algorithm is also implemented to examine US public opinion on homeownership as the best long-term investment following the Great Recession. The results provide interesting insights which may be useful for policymakers and researchers on US housing market.

The GAL distribution proposed by Yan and Kottas (2017) and further studied in this chapter is relatively new and needs to be studied further, particularly due to its usefulness in Bayesian quantile regression. In fact, the GAL distribution can practically be employed to estimate most Bayesian quantile regression models that have been estimated using the AL distribution. A partial list includes the Tobit model with endogenous covariates, censored model, count data model, mixed-effect or longitudinal data model (work in progress), and censored dynamic panel data model. Moreover, the distribution can also be utilized to explore Bayesian variable selection in all the above-mentioned models. We leave these opportunities for future research.

ACKNOWLEDGMENTS

We thank Professor Ivan Jeliazkov (University of California, Irvine) and the anonymous referee for some valuable comments. Discussion and suggestions from the participants at the Winter School, Delhi School of Economics (2016, New Delhi, India), Asian Meeting of the Econometric Society (2017, Hong Kong), Australasian Meeting of the Econometric Society (2018, Auckland, New Zealand), and Joint Statistical Meeting (2018, Vancouver, Canada) are appreciated.

Shubham Karnawat worked on this chapter while pursuing his MS degree in the Department of Economic Sciences at the Indian Institute of Technology Kanpur, India.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transaction on Automatic Control*, *19*(6), 716–723.

Alhamzawi, R., & Ali, H. T. M. (2018). Bayesian quantile regression for ordinal longitudinal data. *Journal of Applied Statistics*, *45*(5), 815–828. doi:10.1080/02664763.2017.1315059

Belsky, E. S. (2013). *The dream lives on: The future of homeownership in America* (Techical Report). Working Paper W11-4 ed. Joint Center for Housing Studies, January 2013, Cambridge, MA.

Benoit, D. F., & Poel, D. V. D. (2010). Binary quantile regression: A Bayesian approach based on the asymmetric Laplace distribution. *Journal of Applied Econometrics*, *27*(7), 1174–1188.

Botev, Z. (2017). The normal law under linear restrictions: Simulation and estimation via minimax tilting. *Journal of the Royal Statistical Society – Series B*, *79*(1), 125–148. doi:10.1111/rssb.12162

Bracha, A., & Jamison, J. C. (2012). Shifting confidence in homeownership: The Great Recession. *B.E. Journal of Macroeconomics*, *12*(3), 1–46. doi:10.1515/1935-1690.104

Dagpunar, J. (1988). *Principles of random variate generation*. Oxford: Clarendon Press.

Dagpunar, J. (1989). An easily implemented generalized inverse Gaussian generator. *Communications in Statistics – Simulation and Computation*, *18*(2), 703–710.

Dagpunar, J. (2007). *Simulations and Monte Carlo: With applications in finance and MCMC*. Chichester: John Wiley & Sons Ltd.

Devroye, L. (2014). Random variate generation for the generalized inverse Gaussian distribution. *Statistics and Computing*, *24*(2), 239–246.

Drew, R. B., & Herbert, C. E. (2013). Postrecession drivers of preferences for homeownership. *Housing Policy Debate*, *23*(4), 666–687.

Geraci, M., & Bottai, M. (2007). Quantile regression for longitudinal data using the asymmetric Laplace distribution. *Biostatistics*, *8*(1), 140–154.

Giuliano, P., & Spilimbergo, A. (2014). Growing up in a recession. *The Review of Economic Studies*, *81*(2), 787–817. doi:10.1093/restud/rdt040

Greenberg, E. (2012). *Introduction to Bayesian econometrics*. New York, NY: Cambridge University Press.

Jeliazkov, I., Graves, J., & Kutzbach, M. (2008). Fitting and comparison of models for multivariate ordinal outcomes. *Advances in Econometrics: Bayesian Econometrics*, *23*, 115–156.

Jeliazkov, I., & Rahman, M. A. (2012). Binary and ordinal data analysis in economics: Modeling and estimation. In X. S. Yang (Ed.), *Mathematical modeling with multidisciplinary applications* (pp. 123–150). Hoboken, NJ: John Wiley & Sons Inc.

Jeliazkov, I., & Vossmeyer, A. (2018). The impact of estimation uncertainty on covariate effects in nonlinear models. *Statistical Papers*, *59*(3), 1031–1042.

Johnson, V. E., & Albert, J. H. (2000). *Ordinal data modeling*. New York, NY: Springer.

Kobayashi, G. (2017). Bayesian endogenous Tobit quantile regression. *Bayesian Analysis*, *12*(1), 161–191. doi:10.1214/16-BA996

Koenker, R. (2005). *Quantile regression*. Cambridge: Cambridge University Press.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica*, *46*(1), 33–50.

Koenker, R., & Machado, J. A. F. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, *94*(448), 1296–1310.

Kotz, S., Kozubowski, T. J., & Podgorski, K. (2001). *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering and finance*. Boston, MA: Birkhäuser.

Kozumi, H., & Kobayashi, G. (2011). Gibbs sampling methods for Bayesian quantile regression. *Journal of Statistical Computation and Simulation*, *81*(11), 1565–1578.

Kozumi, H., & Kobayashi, G. (2012). Bayesian analysis of quantile regression for censored dynamic panel data model. *Computational Statistics*, *27*(2), 359–380.

Lee, D., & Neocleous, T. (2010). Bayesian quantile regression for count data with application to environmental epidemiology. *Journal of the Royal Statistical Society – Series C*, *59*(5), 905–920.

Luo, Y., Lian, H., & Tian, M. (2012). Bayesian quantile regression for longitudinal data models. *Journal of Statistical Computation and Simulation*, *82*(11), 1635–1649.

Malmendier, U., & Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics*, *126*(1), 373–416. doi:10.1093/qje/qjq004

Miguéis, V. L., Benoit, D. F., & Poel, D. V. D. (2013). Enhanced decision support in credit scoring using Bayesian binary quantile regression. *Journal of the Operational Research Society*, *64*(9), 1374–1383.

Mukherjee, D., & Rahman, M. A. (2016). To drill or not to drill? an econometric analysis of us public opinion. *Energy Policy*, *91*, 341–351. doi:10.1016/j.enpol.2015.11.023

Omata, Y., Katayama, H., & Arimura, T. H. (2017). Same concerns, same responses? A Bayesian regression analysis of the determinants for supporting nuclear power generation in Japan. *Environmental Economics and Policy Studies*, *19*(3), 581–608.

Rahman, M. A. (2016). Bayesian quantile regression for ordinal models. *Bayesian Analysis*, *11*(1), 1–24. doi:10.1214/15-BA939

Rahman, M. A., & Vossmeyer, A. (2019). Estimation and applications of quantile regression for binary longitudinal data. *Advances in Econometrics*, *40B*, 157–191.

Reed, C., & Yu, K. (2009). *A partially collapsed Gibbs sampler for Bayesian quantile regression*. Technical Report. Department of Mathematical Sciences, Brunel University, Uxbridge. Retrieved from http://bura.brunel.ac.uk/handle/2438/3593

Reich, B. J., & Smith, L. B. (2013). Bayesian quantile regression for censored data. *Biometrics*, *69*(3), 651–660.

Rohe, W. M., & Lindblad, M. (2014). Reexamining the social benefits of homeownership after the housing crisis. In E. S. Belsky, C. E. Herbert, & J. H. Molinsky (Eds.), *Homeownership built*

to last: Balancing access, affordability, and risk after the housing crisis (pp. 99–140). Washington, DC: Brookings Institution Press.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464.

Tsionas, E. (2003). Bayesian quantile inference. *Journal of Statistical Computation and Simulation*, *73*(9), 659–674.

van Dyk, D. A., & Jiao, X. (2015). Metropolis-Hastings within partially collapsed Gibbs samplers. *Journal of Computational and Graphical Statistics*, *24*(2), 301–327.

Yan, Y., & Kottas, A. (2017). *A new family of error distributions for Bayesian quantile regression*. Working Paper. University of California Santa Cruz.

Yu, K., & Moyeed, R. A. (2001). Bayesian quantile regression. *Statistics and Probability Letters*, *54*(4), 437–447.

Yu, K., & Stander, J. (2007). Bayesian analysis of a Tobit quantile regression model. *Journal of Econometrics*, *137*(1), 260–276.

Yu, K., & Zhang, J. (2005). A three paramter asymmetric Laplace distribution and its extensions. *Communications in Statistics – Theory and Methods*, *34*(9–10), 1867–1879.

Zhou, L. (2010). *Conditional quantile estimation with ordinal data.* Unpublished doctoral dissertation, University of South Carolina.

AUTHOR BIOGRAPHY

Mohammad Arshad Rahman is an Assistant Professor in the Department of Economic Sciences at the Indian Institute of Technology Kanpur, India. He obtained his PhD in Economics from the University of California, Irvine, and holds three Master's degrees. His research interests include Bayesian Econometrics, Quantile Regression, Discrete Choice Models, Markov Chain Monte Carlo (MCMC) methods, Computation and Data Analysis, and Applied Econometrics.

Shubham Karnawat is a Risk Analyst at Credit Suisse, working with the model risk management team. Shubham received his BS and MS in Economics with distinction from the Indian Institute of Technology Kanpur, India. His research interests include Bayesian Econometrics, Quantile Regression, and Machine Learning in Economics.

APPENDIX 1: THE GAL DISTRIBUTION

This appendix derives the *pdf* of the GAL distribution from the mixture representation of the AL distribution and introduces the *cdf* and *mgf* of the GAL distribution. The *mgf* is also utilized to derive the mean, variance, and skewness of the GAL distribution.

A.1. PROBABILITY DENSITY FUNCTION

Theorem 1. Suppose $Y \sim GAL(\mu, \sigma, p, \alpha)$ and has the *pdf* given by Eq. (2), then Y has the following hierarchical representation, $Y = \mu + \alpha\sigma S + \sigma A(p)W + \sigma[B(p)W]^{\frac{1}{2}}U$, where all the notations are as in Section 2.

Proof. Using the mixture representation we can write the *pdf* of Y as

$$\begin{aligned} f(y|\theta) &= \int_{\mathbb{R}^+}\int_{\mathbb{R}^+} N(y|\mu+\sigma\alpha s+\sigma A(p)w, \sigma^2 B(p)w)\exp(w|1)N^+(s|0,1)dw\,ds \\ &= \int_{\mathbb{R}^+} \frac{1}{\sqrt{\sigma^2 B(p)}}\int_0^\infty \left[\frac{1}{\sqrt{2\pi w}}\exp\left\{-\frac{1}{2}\left[\frac{(y-b-aw)^2}{cw}\right]-w\right\}dw\right] \\ &\quad N^+(s|0,1)\,ds \end{aligned} \tag{A.1}$$

where $\boldsymbol{\theta} = (\mu, \sigma, p, \alpha)$, $a = \sigma A(p) = \frac{\sigma(1-2p)}{p(1-p)}$, $b = \mu + \sigma\alpha s$, and $c = \sigma^2 B(p) = \frac{2\sigma^2}{p(1-p)}$. We let P denote the second integral and integrate with respect to w as follows,

$$\begin{aligned} P &= \int_0^\infty \frac{1}{\sqrt{2\pi w}}\exp\left\{-\frac{1}{2}\left[\frac{(a^2+2c)w^2+(y-b)^2-2a(y-b)w)}{cw}\right]\right\}dw \\ &= \exp\left\{\frac{a(y-b)}{c}\right\}\int_0^\infty \frac{1}{\sqrt{2\pi w}}\exp\left\{-\frac{1}{2}\left[\frac{(a^2+2c)}{c}w+\frac{(y-b)^2}{cw}\right]\right\}dw \\ &= \exp\left\{\frac{a(y-b)}{c}\right\}\int_0^\infty \sqrt{\frac{\gamma}{2\pi w}}\sqrt{\frac{1}{\gamma}}\exp\left\{-\frac{1}{2}\left[\gamma w+\frac{\gamma}{\mu^2 w}-\frac{2\gamma}{\mu}+\frac{2\gamma}{\mu}\right]\right\}dw \\ &= \frac{1}{\sqrt{\gamma}}\exp\left\{-\frac{\gamma}{\mu}\right\}\exp\left\{\frac{a(y-b)}{c}\right\}\underbrace{\int_0^\infty \sqrt{\frac{\gamma}{2\pi w}}\exp\left\{-\frac{\gamma}{2}\left(\frac{(1-\mu w)^2}{\mu^2 w}\right)\right\}dw}_{\text{integrates to 1}} \\ &= \frac{1}{\sqrt{\gamma}}\exp\left\{-\frac{\gamma}{\mu}\right\}\exp\left\{\frac{a(y-b)}{c}\right\} \quad (\text{for } \gamma, \mu > 0), \end{aligned} \tag{A.2}$$

where the third line makes the substitutions $\gamma = \frac{a^2+2c}{c} = \frac{1}{2p(1-p)}$, $\mu^2 = \frac{\gamma c}{(y-b)^2} = \frac{\sigma^2}{p^2(1-p)^2(y-b)^2}$, and $\mu = \frac{\sigma}{p(1-p)|y-b|}$. In the fourth line, integration with respect to w yields 1 because it is the *pdf* of a reciprocal inverse-Gaussian

distribution, that is, $w \sim RIG(\gamma, \mu)$. Substituting the values of (γ, μ, a, b, c) in Eq. (A.2) and canceling terms, we get

$$\begin{aligned} P &= \sqrt{2p(1-p)}\ \exp\left\{\frac{(1-2p)(y-b)}{2\sigma} - \frac{|y-b|}{2\sigma}\right\} \\ &= \sqrt{2p(1-p)}\ \exp\left\{-\frac{1}{\sigma}[p - I(y \le b)](y-b)\right\} \\ &= \sqrt{2p(1-p)}\ \exp\left\{-\frac{1}{\sigma}[p - I(y \le \mu + \sigma\alpha s)](y-\mu-\sigma\alpha s)\right\}. \end{aligned} \tag{A.3}$$

Substituting the value of P from Eq. (A.3) into Eq. (A.1), canceling terms, writing the *pdf* of S and letting $\kappa = 2p(1-p)/\sigma$, the *pdf* of Y is

$$\begin{aligned} f(y|\theta) &= \kappa \int_{\mathbb{R}^+} \exp\left\{-\frac{1}{\sigma}[p - I(y \le \mu + \sigma\alpha s)](y-\mu-\sigma\alpha s)\right\} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{s^2}{2}\right\} ds \\ &= \kappa \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[s^2 + 2\left(\frac{y-\mu}{\sigma} - \alpha s\right)\left[p - I\left(\frac{y-\mu}{\sigma} \le \alpha s\right)\right]\right]\right\} ds. \end{aligned} \tag{A.4}$$

Evaluation of the *pdf* $f(y|\theta)$ given by Eq. (A.4) leads to four cases depending on the sign of α and $y^* = (y-\mu)/\sigma$, and we integrate them one at a time. We also employ the earlier introduced notation $p_{\alpha_-} = p - I(\alpha < 0)$ and $p_{\alpha_+} = p - I(\alpha > 0)$ in each cases.

Case (i). When $(\alpha > 0,\ y^* \le 0)$, then $I(y^* \le \alpha s) = 1$. The corresponding *pdf* is,

$$\begin{aligned} f(y|\theta) &= \kappa \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[s^2 + 2\left(\frac{y-\mu}{\sigma} - \alpha s\right)(p-1)\right]\right\} ds \\ &= \kappa \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[s^2 + 2(y^* - \alpha s)p_{\alpha_+}\right]\right\} ds \\ &= \kappa \exp\left\{-y^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\} \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(s - \alpha p_{\alpha_+}\right)^2\right\} ds \\ &= \kappa \exp\left\{-y^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\} \Phi\left(s - \alpha p_{\alpha_+}\right)\Big|_0^\infty \\ &= \kappa\, \Phi\left(\alpha p_{\alpha_+}\right) \exp\left\{-y^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}. \end{aligned} \tag{A.5}$$

Case (ii). For ($\alpha>0$, $y^*>0$) we have two cases. Case (a): $y^*>\alpha s$ implies $I(y^*\le\alpha s)=0$ and this occurs for all $s\in[0, y^*/\alpha)$. Case(b): $y^*\le\alpha s$ implies $I(y^*\le\alpha s)=1$, and this occurs for all $s\in[y^*/\alpha,\infty)$. Hence, the *pdf* is

$$
\begin{aligned}
f(y\,|\,\theta) &= \kappa\int_0^{\frac{y^*}{\alpha}}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[s^2+2(y^*-\alpha s)p\right]\right\}ds \\
&\quad+\kappa\int_{\frac{y^*}{\alpha}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[s^2+2(y^*-\alpha s)(p-1)\right]\right\}ds \\
&=\kappa\exp\left\{-y^*p_{\alpha_-}\right\}\int_0^{\frac{y^*}{\alpha}}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(s^2-2\alpha p_{\alpha_-}s\right)\right\}ds \\
&\quad+\kappa\exp\left\{-y^*p_{\alpha_+}\right\}\int_{\frac{y^*}{\alpha}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(s^2-2\alpha p_{\alpha_+}s\right)\right\}ds \\
&=\kappa\exp\left\{-y^*p_{\alpha_-}+\frac{1}{2}\alpha^2p_{\alpha_-}^2\right\}\int_0^{\frac{y^*}{\alpha}}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(s-\alpha p_{\alpha_-}\right)^2\right\}ds \\
&\quad+\kappa\exp\left\{-y^*p_{\alpha_+}+\frac{1}{2}\alpha^2p_{\alpha_+}^2\right\}\int_{\frac{y^*}{\alpha}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(s^2-\alpha p_{\alpha_+}\right)^2\right\}ds \\
&=\kappa\exp\left\{-y^*p_{\alpha_-}+\frac{1}{2}\alpha^2p_{\alpha_-}^2\right\}\Phi\left(s-\alpha p_{\alpha_-}\right)\Big|_0^{\frac{y^*}{\alpha}} \\
&\quad+\kappa\exp\left\{-y^*p_{\alpha_+}+\frac{1}{2}\alpha^2p_{\alpha_+}^2\right\}\Phi\left(s-\alpha p_{\alpha_+}\right)\Big|_{\frac{y^*}{\alpha}}^{\infty} \\
&=\kappa\left[\Phi\left(\frac{y^*}{\alpha}-\alpha p_{\alpha_-}\right)-\Phi(-\alpha p_{\alpha_-})\right]\exp\left\{-y^*p_{\alpha_-}+\frac{1}{2}\alpha^2p_{\alpha_-}^2\right\} \\
&\quad+\kappa\left[\Phi\left(\alpha p_{\alpha_+}-\frac{y^*}{\alpha}\right)\right]\exp\left\{-y^*p_{\alpha_+}+\frac{1}{2}\alpha^2p_{\alpha_+}^2\right\}. \qquad \text{(A.6)}
\end{aligned}
$$

Case (iii). When ($\alpha<0$, $y^*>0$), then $I(y^*\le\alpha s)=0$ since $\alpha<0$. Hence, we have

$$
\begin{aligned}
f(y\,|\,\theta) &= \kappa \int_0^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left[s^2 + 2(y^* - \alpha s) p_{\alpha_+} \right] \right\} ds \\
&= \kappa \exp\left\{ -y^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2 \right\} \int_0^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left(s - \alpha p_{\alpha_+} \right)^2 \right\} ds \\
&= \kappa\, \Phi(\alpha p_{\alpha_+}) \exp\left\{ -y^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2 \right\}.
\end{aligned}
\tag{A.7}
$$

Case (iv). For ($\alpha < 0, y^* \le 0$), we have two cases. Case (a): $y^* \le \alpha s$ implies $I(y^* \le \alpha s) = 1$, and this occurs for all $s \in [0,\ y^*/\alpha]$. Case(b): $y^* > \alpha s$ implies $I(y^* \le \alpha s) = 0$, and this occurs for all $s \in (y^*/\alpha, \infty)$. So we have,

$$
\begin{aligned}
f(y\,|\,\theta) &= \kappa \int_0^{\frac{y^*}{\alpha}} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left[s^2 + 2(y^* - \alpha s)(p-1) \right] \right\} ds \\
&\quad + \kappa \int_{\frac{y^*}{\alpha}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left[s^2 + 2(y^* - \alpha s) p \right] \right\} ds \\
&= \kappa \int_0^{\frac{y^*}{\alpha}} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left[s^2 + 2(y^* - \alpha s) p_{\alpha_-} \right] \right\} ds \\
&\quad + \kappa \int_{\frac{y^*}{\alpha}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left[s^2 + 2(y^* - \alpha s) p_{\alpha_+} \right] \right\} ds \\
&= \kappa \left[\Phi\left(\frac{y^*}{\alpha} - \alpha p_{\alpha_-} \right) - \Phi(-\alpha p_{\alpha_-}) \right] \exp\left\{ -y^* p_{\alpha_-} + \frac{1}{2}\alpha^2 p_{\alpha_-}^2 \right\} \\
&\quad + \kappa \left[\Phi\left(\alpha p_{\alpha_+} - \frac{y^*}{\alpha} \right) \right] \exp\left\{ -y^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2 \right\},
\end{aligned}
\tag{A.8}
$$

where the integration details are similar to Case (ii) and have been suppressed to avoid monotonicity and save space.

Combining all the four cases, that is, Eqs. (A.5) to (A.8), we have the *pdf* of the GAL distribution given by Eq. (2).

A.2. CUMULATIVE DISTRIBUTION FUNCTION

Theorem 2. Suppose $Y \sim GAL(\mu, \sigma, p, \alpha)$ and let $y^* = (y-\mu)/\sigma$, then the *cdf* F is

$$\begin{aligned} F(y\,|\,\theta) = & \left(1 - 2\Phi\left(-\frac{y^*}{|\alpha|}\right) - \frac{2p(1-p)}{p_{\alpha_-}}\exp\left\{-y^* p_{\alpha_-} + \frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}\right. \\ & \left.\left[\Phi\left(\frac{y^*}{\alpha} - \alpha p_{\alpha_-}\right) - \Phi(-\alpha p_{\alpha_-})\right]\right) I\left(\frac{y^*}{\alpha} > 0\right) \\ & + I(\alpha<0) - \frac{2p(1-p)}{p_{\alpha_+}}\exp\left\{-y^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\} \\ & \times \Phi\left[\alpha p_{\alpha_+} - \frac{y^*}{\alpha} I\left(\frac{y^*}{\alpha} > 0\right)\right]. \end{aligned} \tag{A.9}$$

Proof. We note that for any *cdf* $F(y\,|\,\theta) = \int_{-\infty}^{y} f(v\,|\,\theta)dv = 1 - \int_{y}^{\infty} f(v\,|\,\theta)dv$. This property is used in deriving the *cdf* when $y > \mu$ to avoid breaking the region of integration as $(-\infty, \mu) \cup (\mu, y)$. We let $v^* = (v-\mu)/\sigma$, combining cases and derive as follows:

Case (i). When ($\alpha > 0,\ y \le \mu$) or ($\alpha < 0,\ y > \mu$), the *cdf* is

$$\begin{aligned} F(y|\theta) &= \begin{cases} \displaystyle\int_{-\infty}^{y} \kappa\Phi(\alpha p_{\alpha_+})\exp\left\{-v^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}dv, & \text{if } \alpha>0, y\le\mu \\ \displaystyle 1 - \int_{y}^{\infty} \kappa\Phi(\alpha p_{\alpha_+})\exp\left\{-v^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}dv, & \text{if } \alpha<0, y>\mu \end{cases} \\ &= \begin{cases} \kappa\Phi(\alpha p_{\alpha_+})\exp\left\{\frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}\left[\frac{\exp\{-v^* p_{\alpha_+}\}}{-p_{\alpha_+}/\sigma}\right]_{-\infty}^{y}, & \text{if } \alpha>0, y\le\mu \\ 1 - \kappa\Phi(\alpha p_{\alpha_+})\exp\left\{\frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}\left[\frac{\exp\{-v^* p_{\alpha_+}\}}{-p_{\alpha_+}/\sigma}\right]_{y}^{\infty}, & \text{if } \alpha<0, y>\mu \end{cases} \\ &= \begin{cases} \frac{2p(1-p)}{-p_{\alpha_+}}\Phi(\alpha p_{\alpha_+})\exp\left\{-y^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}, & \text{if } \alpha>0, y\le\mu \\ 1 - \frac{2p(1-p)}{p_{\alpha_+}}\Phi(\alpha p_{\alpha_+})\exp\left\{-y^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}, & \text{if } \alpha<0, y>\mu \end{cases} \end{aligned} \tag{A.10}$$

where the third step substitutes the value of κ, $p_{\alpha_+} = p - 1$ for $\alpha > 0$ and $p_{\alpha_+} = p$ for $\alpha < 0$.

Case (ii). When ($\alpha < 0, y \leq \mu$) or ($\alpha > 0, y > \mu$), the *cdf* is,

$$F(y|\theta) = \begin{cases} \int_{-\infty}^{y} \kappa \left(\left[\Phi\left(\frac{v^*}{\alpha} - \alpha p_{\alpha_-} \right) - \Phi(-\alpha p_{\alpha_-}) \right] \exp\left\{ -v^* p_{\alpha_-} + \frac{1}{2}\alpha^2 p_{\alpha_-}^2 \right\} \right. \\ \left. + \Phi\left(\alpha p_{\alpha_+} - \frac{v^*}{\alpha} \right) \exp\left\{ -v^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2 \right\} \right) dv, \quad \text{if } \alpha < 0, \mathrm{y} \leq \mu \\ 1 - \int_{y}^{\infty} \kappa \left(\left[\Phi\left(\frac{v^*}{\alpha} - \alpha p_{\alpha_-} \right) - \Phi(-\alpha p_{\alpha_-}) \right] \exp\left\{ -v^* p_{\alpha_-} + \frac{1}{2}\alpha^2 p_{\alpha_-}^2 \right\} \right. \\ \left. + \Phi\left(\alpha p_{\alpha_+} - \frac{v^*}{\alpha} \right) \exp\left\{ -v^* p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2 \right\} \right) dv, \quad \text{if } \alpha > 0, \mathrm{y} > \mu. \end{cases} \tag{A.11}$$

Note that in both the subcases of Eq. (A.11), the integral remains the same and only the limits of integration change. Hence, we evaluate each terms individually over the limits (a, b) and will substitute values of (a, b) as per our requirement.

To evaluate the first integral component of Eq. (A.11) denoted C_1, we substitute $z = v^*/\alpha - \alpha p_{\alpha_-}$ and perform integration by parts as follows:

$$\begin{aligned} C_1 &= \kappa \int_a^b \Phi\left(\frac{v^*}{\alpha} - \alpha p_{\alpha_-} \right) \exp\left\{ -v^* p_{\alpha_-} + \frac{1}{2}\alpha^2 p_{\alpha_-}^2 \right\} dv \\ &= \kappa \exp\left\{ -\frac{1}{2}\alpha^2 p_{\alpha_-}^2 \right\} \int_{\frac{a^*}{\alpha} - \alpha p_{\alpha_-}}^{\frac{b^*}{\alpha} - \alpha p_{\alpha_-}} \left[\alpha\sigma\Phi(z)\exp\{-\alpha p_{\alpha_-} z\} \right] dz \\ &= \kappa\sigma \exp\left\{ -\frac{1}{2}\alpha^2 p_{\alpha_-}^2 \right\} \left(\left[\frac{\alpha\Phi(z)\exp\{-\alpha p_{\alpha_-} z\}}{-\alpha p_{\alpha_-}} \right]_{\frac{a^*}{\alpha} - \alpha p_{\alpha_-}}^{\frac{b^*}{\alpha} - \alpha p_{\alpha_-}} \right. \\ &\left. - \frac{\alpha}{-\alpha p_{\alpha_-}} \int_{\frac{a^*}{\alpha} - \alpha p_{\alpha_-}}^{\frac{b^*}{\alpha} - \alpha p_{\alpha_-}} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}(z^2 + 2\alpha p_{\alpha_-} z + \alpha^2 p_{\alpha_-}^2 - \alpha^2 p_{\alpha_-}^2) \right\} dz \right) \end{aligned}$$

$$
\begin{aligned}
&= \kappa\sigma \exp\left\{-\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}\left(\frac{\exp\left\{\alpha^2 p_{\alpha_-}^2\right\}}{-p_{\alpha_-}}\left[\Phi\left(\frac{b^*}{\alpha}-\alpha p_{\alpha_-}\right)\exp\{-b^* p_{\alpha_-}\}\right.\right.\\
&\left.\left.-\Phi\left(\frac{a^*}{\alpha}-\alpha p_{\alpha_-}\right)\exp\{-a^* p_{\alpha_-}\}\right]+\frac{1}{p_{\alpha_-}}\exp\left\{\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}\Phi\left(z+\alpha p_{\alpha_-}\right)_{\frac{a^*}{\alpha}-\alpha p_{\alpha_-}}^{\frac{b^*}{\alpha}-\alpha p_{\alpha_-}}\right)\\
&= -\frac{\kappa\sigma}{p_{\alpha_-}}\exp\left\{\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}\left[\Phi\left(\frac{b^*}{\alpha}-\alpha p_{\alpha_-}\right)\exp\{-b^* p_{\alpha_-}\}\right.\\
&\left.-\Phi\left(\frac{a^*}{\alpha}-\alpha p_{\alpha_-}\right)\exp\{-a^* p_{\alpha_-}\}\right]+\frac{\kappa\sigma}{p_{\alpha_-}}\left[\Phi\left(\frac{b^*}{\alpha}\right)-\Phi\left(\frac{a^*}{\alpha}\right)\right].
\end{aligned}
\tag{A.12}
$$

We next evaluate the second component denoted C_2 directly as follows:

$$
\begin{aligned}
C_2 &= -\kappa\Phi(-\alpha p_{\alpha_-})\int_a^b \exp\left\{-v^* p_{\alpha_-}+\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}dv\\
&= -\kappa\Phi(-\alpha p_{\alpha_-})\exp\left\{\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}\left[\frac{\exp\{-v^* p_{\alpha_-}\}}{-p_{\alpha_-}/\sigma}\right]_a^b\\
&= \frac{\kappa\sigma}{p_{\alpha_-}}\Phi(-\alpha p_{\alpha_-})\exp\left\{\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}\left[\exp\{-b^* p_{\alpha_-}\}-\exp\{-a^* p_{\alpha_-}\}\right].
\end{aligned}
\tag{A.13}
$$

Finally, to evaluate the third component denoted C_3 we use the substitution $z = \alpha p_{\alpha_+} - v^*/\alpha$ and integrate-by-parts as done in C_1. We suppress the details for brevity and present the final expression:

$$
\begin{aligned}
C_3 &= \kappa\int_a^b \Phi\left(\alpha p_{\alpha_+}-\frac{v^*}{\alpha}\exp\left\{-v^* p_{\alpha_+}+\frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}\right)dv\\
&= -\frac{\kappa\sigma}{p_{\alpha_+}}\exp\left\{\frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}\left[\Phi\left(\alpha p_{\alpha_+}-\frac{b^*}{\alpha}\right)\right.\\
&\left.\exp\{-b^* p_{\alpha_+}\}-\Phi\left(\alpha p_{\alpha_+}-\frac{a^*}{\alpha}\right)\exp\{-a^* p_{\alpha_+}\}\right]\\
&+\frac{\kappa\sigma}{p_{\alpha_+}}\left[\Phi\left(-\frac{b^*}{\alpha}\right)-\Phi\left(-\frac{a^*}{\alpha}\right)\right].
\end{aligned}
\tag{A.14}
$$

Note that $\Phi\left(\frac{b^*}{\alpha}\right) - \Phi\left(\frac{a^*}{\alpha}\right) = -\Phi\left(-\frac{b^*}{\alpha}\right) + \Phi\left(-\frac{a^*}{\alpha}\right)$, and hence, the relevant term from Eqs. (A.12) and (A.14) can be collected together when adding the expressions.

When $\alpha<0$ and $y\leq\mu$, the limits of integration $a=-\infty$ and $b=y$ imply $a^*=-\infty$ and $b^*=y^*$, respectively. Substituting the values of a^*, b^* and κ in C_1, C_2 and C_3 and summing the expression yields,

$$\begin{aligned} F(y\,|\,\theta) \;=\; & 2\Phi\left(-\frac{y^*}{\alpha}\right) - \frac{2p(1-p)}{p_{\alpha_-}}\exp\left\{-y^*p_{\alpha_-} + \frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\} \\ & \left[\Phi\left(\frac{y^*}{\alpha} - \alpha p_{\alpha_-}\right) - \Phi(-\alpha p_{\alpha_-})\right] \\ & - \frac{2p(1-p)}{p_{\alpha_+}}\exp\left\{-y^*p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}\Phi\left(\alpha p_{\alpha_+} - \frac{y^*}{\alpha}\right). \end{aligned} \tag{A.15}$$

Similarly, when $\alpha>0$ and $y>\mu$, the limits of integration $a=y$ and $b=\infty$ imply $a^*=y^*$ and $b^*=\infty$, respectively. Substituting the values of a^*, b^* and κ in C_1, C_2 and C_3 and evaluating the expression $1-C_1-C_2-C_3$ yield

$$\begin{aligned} F(y\,|\,\theta) = 1 - & 2\Phi\left(-\frac{y^*}{\alpha}\right) - \frac{2p(1-p)}{p_{\alpha_-}}\exp\left\{-y^*p_{\alpha_-} + \frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\} \\ & \left[\Phi\left(\frac{y^*}{\alpha} - \alpha p_{\alpha_-}\right) - \Phi(-\alpha p_{\alpha_-})\right] - \frac{2p(1-p)}{p_{\alpha_+}} \\ & \exp\left\{-y^*p_{\alpha_+} + \frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}\Phi\left(\alpha p_{\alpha_+} - \frac{y^*}{\alpha}\right). \end{aligned} \tag{A.16}$$

Combining the Eqs. (A.10), (A.15), and (A.16), we have the *cdf* of the GAL distribution given by Eq. (3).

A.3. MOMENT GENERATING FUNCTION

Theorem 3. Suppose $Y\sim GAL(\mu,\sigma,p,\alpha)$, then the *mgf* denoted $M_Y(t)$ is as follows,

$$M_Y(t) = 2p(1-p)\left[\frac{(p_{\alpha_+}-p_{\alpha_-})}{(p_{\alpha_-}-\sigma t)(p_{\alpha_+}-\sigma t)}\right]\exp\left\{\mu t + \frac{1}{2}\alpha^2\sigma^2 t^2\right\}\Phi(|\,\alpha\,|\,\sigma t). \tag{A.17}$$

Proof. Using the definition of the *mgf*, we have

$$M_Y(t) = \int_{-\infty}^{\infty} \exp(ty) f(y \mid \mu, \sigma, p, \alpha)\, dy. \tag{A.18}$$

Substituting the GAL *pdf* given by Eq. (2) into Eq. (A.18) leads to two cases depending on $\alpha > 0$ or $\alpha < 0$. We again use the notation $\kappa = 2p(1-p)/\sigma$ and break the region of integration depending on $y^* > 0$ (i.e., $y > \mu$) or $y^* \leq 0$ (i.e., $y \leq \mu$).

Case (i). When $\alpha > 0$ and $y^* > 0$ (i.e., $y > \mu$), we have the following three components,

$$\begin{aligned}
M_1 &= \kappa \int_{\mu}^{\infty} \exp\{ty\}\, \Phi\left(\frac{y^*}{\alpha} - \alpha p_{\alpha_-}\right) \exp\left\{-y^* p_{\alpha_-} + \frac{1}{2}(\alpha p_{\alpha_-})^2\right\} dy, \\
M_2 &= -\kappa \int_{\mu}^{\infty} \exp\{ty\}\, \Phi(-\alpha p_{\alpha_-}) \exp\left\{-y^* p_{\alpha_-} + \frac{1}{2}(\alpha p_{\alpha_-})^2\right\} dy, \\
M_3 &= \kappa \int_{\mu}^{\infty} \exp\{ty\}\, \Phi\left(\alpha p_{\alpha_+} - \frac{y^*}{\alpha}\right) \exp\left\{-y^* p_{\alpha_+} + \frac{1}{2}(\alpha p_{\alpha_+})^2\right\} dy,
\end{aligned} \tag{A.19}$$

and when $\alpha > 0$ and $y^* \leq 0$ (i.e., $y \leq \mu$), we have

$$M_4 = \kappa \int_{-\infty}^{\mu} \exp\{ty\}\, \Phi(\alpha p_{\alpha_+}) \exp\left\{-y^* p_{\alpha_+} + \frac{1}{2}(\alpha p_{\alpha_+})^2\right\} dy. \tag{A.20}$$

We first consider M_1, substitute $z = \frac{y^*}{\alpha} - \alpha p_{\alpha_-}$, change the limits of integration, and integrate-by-parts as follows:

$$\begin{aligned}
M_1 &= \int_{-\alpha p_{\alpha_-}}^{\infty} \kappa\alpha\sigma \exp\{\mu t + \alpha\sigma(z + \alpha p_{\alpha_-})t\}\, \Phi(z) \exp\left\{-\alpha p_{\alpha_-}(z + \alpha p_{\alpha_-}) + \frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\} dz \\
&= \kappa\alpha\sigma \exp\left\{\mu t + \alpha^2 p_{\alpha_-}\sigma t - \frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\} \int_{-\alpha p_{\alpha_-}}^{\infty} \Phi(z) \exp\{-\alpha(p_{\alpha_-} - \sigma t)z\} dz \\
&= \kappa\alpha\sigma \exp\left\{\mu t + \alpha^2 p_{\alpha_-}\sigma t - \frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\} \left[\Phi(z) \frac{\exp\{-\alpha(p_{\alpha_-} - \sigma t)z\}}{-\alpha(p_{\alpha_-} - \sigma t)}\Bigg|_{-\alpha p_{\alpha_-}}^{\infty}\right. \\
&\quad \left. - \int_{-\alpha p_{\alpha_-}}^{\infty} \phi(z) \frac{\exp\{-\alpha(p_{\alpha_-} - \sigma t)z\}}{-\alpha(p_{\alpha_-} - \sigma t)} dz\right] \quad (\text{limit exists only if } t < p/\sigma) \\
&= \kappa\sigma \frac{\exp\left\{\mu t + \alpha^2 p_{\alpha_-}\sigma t - \frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}}{(p_{\alpha_-} - \sigma t)} \left[\exp\{\alpha^2 p_{\alpha_-}(p_{\alpha_-} - \sigma t)\}\Phi(-\alpha p_{\alpha_-})\right.
\end{aligned}$$

$$+\exp\left\{\frac{\alpha^2}{2}(p_{\alpha_-}-\sigma t)^2\right\}\int_{-\alpha p_{\alpha_-}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}(z+\alpha(p_{\alpha_-}-\sigma t)^2)\right\}dz\Bigg]$$

$$=\kappa\sigma\frac{\exp\left\{\mu t+\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}}{(p_{\alpha_-}-\sigma t)}\Phi(-\alpha p_{\alpha_-})+\kappa\sigma\frac{\exp\left\{\mu t+\frac{1}{2}\alpha^2\sigma^2 t^2\right\}}{(p_{\alpha_-}-\sigma t)}\Phi(z+\alpha(p_{\alpha_-}-\sigma t))\Bigg|_{-\alpha p_{\alpha_-}}^{\infty}$$

$$=\kappa\sigma\frac{\exp\left\{\mu t+\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}}{(p_{\alpha_-}-\sigma t)}\Phi(-\alpha p_{\alpha_-})+\kappa\sigma\frac{\exp\left\{\mu t+\frac{1}{2}\alpha^2\sigma^2 t^2\right\}}{(p_{\alpha_-}-\sigma t)}\Phi(\alpha\sigma t). \tag{A.21}$$

Second, we integrate the expression for M_2 as follows:

$$M_2=-\kappa\int_{\mu}^{\infty}\exp\{ty\}\Phi(-\alpha p_{\alpha_-})\exp\left\{-y^* p_{\alpha_-}+\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}dy$$

$$=-\kappa\Phi(-\alpha p_{\alpha_-})\exp\left\{\frac{p_{\alpha_-}\mu}{\sigma}+\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}\int_{\mu}^{\infty}\exp\left\{-\frac{(p_{\alpha_-}-\sigma t)}{\sigma}y\right\}dy$$

$$=-\kappa\,\Phi(-\alpha p_{\alpha_-})\exp\left\{\frac{p_{\alpha_-}\mu}{\sigma}+\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}\times\frac{(-1)(-\sigma)}{(p_{\alpha_-}-\sigma t)}\exp\left\{\frac{-(p_{\alpha_-}-\sigma t)}{\sigma}\mu\right\}$$

$$=-\kappa\sigma\frac{\exp\left\{\mu t+\frac{1}{2}\alpha^2 p_{\alpha_-}^2\right\}}{(p_{\alpha_-}-\sigma t)}\Phi(-\alpha p_{\alpha_-})\quad\text{(limit exists only if t}<p/\sigma).$$

(A.22)

The integral M_3 is evaluated by substituting $z=\alpha p_{\alpha_+}-\frac{y^*}{\alpha}$. Thereafter, the integration is analogous to M_1 with the existence condition $t>(p-1)/\alpha$ and yields:

$$M_3=\kappa\sigma\frac{\exp\{\mu t+\frac{1}{2}\alpha^2 p_{\alpha_+}^2\}}{(p_{\alpha_+}-\sigma t)}\Phi(\alpha p_{\alpha_+})-\kappa\sigma\frac{\exp\{\mu t+\frac{1}{2}\alpha^2\sigma^2 t^2\}}{(p_{\alpha_+}-\sigma t)}\Phi(\alpha\sigma t). \tag{A.23}$$

The integral M_4 can be directly evaluated and closely follows the steps in the integration of M_2 with the resulting expression,

$$M_4=-\kappa\sigma\frac{\exp\left\{\mu t+\frac{1}{2}\alpha^2 p_{\alpha_+}^2\right\}}{(p_{\alpha_+}-\sigma t)}\ \Phi(\alpha p_{\alpha_+}). \tag{A.24}$$

where the limits of integration exist only for $t>(p-1)/\sigma$.

To obtain the *mgf* of the GAL distribution, we substitute the values of κ and sum the four Eqs. (A.21)–(A.24). This yields,

$$M_Y(t) = 2p(1-p)\left[\frac{(p_{\alpha_+} - p_{\alpha_-})}{(p_{\alpha_-} - \sigma t)(p_{\alpha_+} - \sigma t)}\right]\exp\left\{\mu t + \frac{1}{2}\alpha^2\sigma^2 t^2\right\}\Phi(\alpha\sigma t). \quad \text{(A.25)}$$

Case (ii). When $\alpha < 0$ and $y^* \leq 0$ (i.e., $y \leq \mu$), we have the following three components of the *mgf*,

$$\begin{aligned}
M_5 &= \kappa \int_{-\infty}^{\mu} \exp\{ty\}\Phi\left(\frac{y^*}{\alpha} - \alpha p_{\alpha_-}\right)\exp\left\{-y^* p_{\alpha_-} + \frac{1}{2}(\alpha p_{\alpha_-})^2\right\}dy, \\
M_6 &= -\kappa \int_{-\infty}^{\mu} \exp\{ty\}\Phi(-\alpha p_{\alpha_-})\exp\left\{-y^* p_{\alpha_-} + \frac{1}{2}(\alpha p_{\alpha_-})^2\right\}dy, \\
M_7 &= \kappa \int_{-\infty}^{\mu} \exp\{ty\}\Phi\left(\alpha p_{\alpha_+} - \frac{y^*}{\alpha}\right)\exp\left\{-y^* p_{\alpha_+} + \frac{1}{2}(\alpha p_{\alpha_+})^2\right\}dy,
\end{aligned} \quad \text{(A.26)}$$

and when $\alpha < 0$ and $y^* > 0$ (i.e., $y > \mu$), we have

$$M_8 = \kappa \int_{\mu}^{\infty} \exp\{ty\}\Phi(\alpha p_{\alpha_+})\exp\left\{-y^* p_{\alpha_+} + \frac{1}{2}(\alpha p_{\alpha_+})^2\right\}dy. \quad \text{(A.27)}$$

The integration of the terms M_5 to M_8 are similar to the case when $\alpha > 0$ and results in the following *mgf*:

$$M_Y(t) = 2p(1-p)\left[\frac{(p_{\alpha_+} - p_{\alpha_-})}{(p_{\alpha_-} - \sigma t)(p_{\alpha_+} - \sigma t)}\right]\exp\left\{\mu t + \frac{1}{2}\alpha^2\sigma^2 t^2\right\}\Phi(-\alpha\sigma t). \quad \text{(A.28)}$$

Combining the *mgf* for the two cases, that is, Eqs. (A.25) and (A.28), we have the *mgf* of the GAL distribution.

The mean, variance, and skewness of the GAL distribution can be obtained from the GAL *mgf* given by Eq. (A.17). We state this in terms of a theorem below.

Theorem 4. Suppose $Y \sim GAL(\mu, \sigma, p, \alpha)$, then

$$\begin{aligned}
E(Y) &= \mu + \frac{2\,|\,\alpha\,|\,\sigma}{\sqrt{2\pi}} + \sigma\frac{(1-2p)}{p(1-p)} \\
V(Y) &= \alpha^2\sigma^2\left(1 - \frac{2}{\pi}\right) + \sigma^2\left[\frac{1-2p+2p^2}{p^2(1-p)^2}\right] \\
S(Y) &= \frac{\alpha^3\sqrt{\frac{2}{\pi}}\left(\frac{4}{\pi} - 1\right) + 2\left[\frac{(1-p)^3 - p^3}{p^3(1-p)^3}\right]}{\left\{\alpha^2\left(1 - \frac{2}{\pi}\right) + \left[\frac{1-2p+2p^2}{p^2(1-p)^2}\right]\right\}^{3/2}}
\end{aligned} \quad \text{(A.29)}$$

where $E(Y)$, $V(Y)$, and $S(Y)$ denote the mean, variance, and skewness, respectively.

Proof. Taking logarithm of the *mgf* and keeping terms involving t, we have

$$\ln M_Y(t) \propto -\ln(p_{\alpha_-} - \sigma t) - \ln(p_{\alpha_+} - \sigma t) + \mu t + \frac{1}{2}\alpha^2\sigma^2 t^2 + \ln \Phi(|\alpha|\sigma t). \tag{A.30}$$

Taking the first, second, and third derivatives of Eq. (A.30) and evaluating at $t = 0$, we get

$$m_1(Y) = E(Y) = \frac{\partial \ln M_Y(t)}{\partial t}\bigg|_{t=0} = \mu + \sqrt{\frac{2}{\pi}}\,\alpha\sigma + \sigma\frac{(1-2p)}{p(1-p)}$$

$$m_2(Y) = V(Y) = \frac{\partial^2 \ln M_Y(t)}{\partial t^2}\bigg|_{t=0} = \alpha^2\sigma^2\left(1 - \frac{2}{\pi}\right) + \sigma^2\left[\frac{1-2p+2p^2}{p^2(1-p)^2}\right]$$

$$m_3(Y) = \frac{\partial^3 \ln M_Y(t)}{\partial t^3}\bigg|_{t=0} = \alpha^3\sigma^3\sqrt{\frac{2}{\pi}}\left(\frac{4}{\pi} - 1\right) + 2\sigma^3\left[\frac{(1-p)^3 - p^3}{p^3(1-p)^3}\right]$$

where $m_1(Y), m_2(Y)$, and $m_3(Y)$ are the first-, second-, and third-order central moments, respectively. Hence, skewness can be obtained as $\frac{m_3}{(m_2^{3/2})}$.

APPENDIX 2: CONDITIONAL DENSITIES IN THE FBQROR MODEL

In this appendix, we derive the conditional posteriors of the FBQROR model parameters. Specifically, the conditional posteriors of β, ν, h, and z have tractable distributions and is sampled using a Gibbs approach. The parameters (σ, γ) are jointly sampled using random-walk MH algorithm (to reduce autocorrelation in MCMC draws) and δ is sampled using a random-walk MH algorithm. The derivations below follow the ordering as presented in Algorithm 1.

(1) Starting with β, the conditional posterior $\pi(\beta \mid z, \nu, h, \sigma, \gamma)$ is proportional to $\pi(\beta) \times f(z \mid \beta, \nu, h, \sigma, \gamma)$ and its kernel can be written as,

$$\begin{aligned}
\pi(\beta|z,\nu,h,\sigma,\gamma) \propto & \exp\left\{ -\frac{1}{2}\left[\sum_{i=1}^{n} \frac{(z_i - x_i'\beta - A\nu_i - C|\gamma|h_i)^2}{\sigma B \nu_i} + (\beta-\beta_0)' B_0^{-1} (\beta-\beta_0) \right] \right\} \\
\propto & \exp\left\{ -\frac{1}{2}\left[\beta' \left(\sum_{i=1}^{n} \frac{x_i x_i'}{\sigma B \nu_i} + B_0^{-1} \right) \beta - \beta' \left(\frac{x_i (z_i - A\nu_i - C|\gamma|h_i)}{\sigma B \nu_i} + B_0^{-1}\beta_0 \right) \right.\right. \\
& \left.\left. - \left(\frac{x_i'(z_i - A\nu_i - C|\gamma|h_i)}{\sigma B \nu_i} + \beta_0' B_0^{-1} \right) \right] \right\} \\
\propto & \exp\left\{ -\frac{1}{2}\left[\beta' \tilde{B}^{-1}\beta - \beta' \tilde{B}^{-1}\tilde{\beta} - \tilde{\beta}' \tilde{B}^{-1}\beta + \tilde{\beta}' \tilde{B}^{-1}\tilde{\beta} - \tilde{\beta}' \tilde{B}^{-1}\tilde{\beta} \right] \right\} \\
\propto & \exp\left\{ -\frac{1}{2}(\beta-\tilde{\beta})' \tilde{B}^{-1} (\beta-\tilde{\beta}) \right\},
\end{aligned}$$

where the posterior variance $\tilde{B}$ and the posterior mean $\tilde{\beta}$ are defined as follows:

$$\tilde{B}^{-1} = \left(\sum_{i=1}^{n} \frac{x_i x_i'}{\sigma B \nu_i} + B_0^{-1} \right) \quad \text{and} \quad \tilde{\beta} = \tilde{B} \left(\sum_{i=1}^{n} \frac{x_i (z_i - A\nu_i - C\,|\,\gamma\,|\,h_i)}{\sigma B \nu_i} + B_0^{-1}\beta_0 \right).$$

Hence, the conditional posterior is a normal distribution and $\beta \mid z, \nu, h, \sigma, \gamma \sim N(\tilde{\beta}, \tilde{B})$.

(2) The parameters (σ, γ) are jointly sampled marginally of (z, ν, h) to reduce autocorrelation in the MCMC draws. Collecting terms involving (σ, γ) from the complete joint posterior density (Eq. 15) does not yield a tractable distribution, hence (σ, γ) are sampled using a joint random-walk MH algorithm. The proposed values are generated from a truncated bivariate normal distribution $TBN_{(0,\infty)\times(L,U)}((\sigma_c, \gamma_c), \iota_1^2 \hat{D}_1)$, where (σ_c, γ_c) represent

the current values, ι_1 denotes the tuning factor, and $\hat{D}_1$ is the negative inverse of the Hessian obtained by maximizing the log-likelihood (Eq. 9) with respect to (σ, γ). This maximization process to obtain $\hat{D}_1$ is computationally expensive and so may be done only once at the beginning of the algorithm with β values fixed at the BQROR or OLS estimates. The proposed draws are accepted with MH probability

$$\alpha_{MH}(\sigma_c,\gamma_c;\sigma',\gamma') = \min\left\{0, \ln\left[\frac{f(y|\beta,\sigma',\gamma',\delta)\pi(\beta,\sigma',\gamma',\delta)}{f(y|\beta,\sigma_c,\gamma_c,\delta)\pi(\beta,\sigma_c,\gamma_c,\delta)}\frac{\pi(\sigma_c,\gamma_c|(\sigma',\gamma'),\iota_1^2\hat{D}_1)}{\pi(\sigma',\gamma'|(\sigma_c,\gamma_c),\iota_1^2\hat{D}_1)}\right]\right\},$$

otherwise, (σ_c, γ_c) is repeated in the next MCMC iteration. Here, $f(\cdot)$ represents the full likelihood (Eq. 9), $\pi(\beta, \sigma, \delta, \gamma)$ denotes the prior distributions (Eq. 13), $\pi(\sigma_c, \gamma_c | (\sigma', \gamma'), \iota_1^2 \hat{D}_1)$ stands for the truncated bivariate normal probability with mean (σ', γ') and covariance $\iota_1^2 \hat{D}_1$. The expression $\pi(\sigma', \gamma' | (\sigma_c, \gamma_c), \iota_1^2 \hat{D}_1)$ has an analogous interpretation. Note that the tuning parameter ι_1 can be adjusted for appropriate step-size and acceptance rate, and the parameters (A, B, C) depend on p which in turn is a function of p_0 and γ.

(3) The conditional posterior of ν is obtained from the complete posterior density (Eq. 15) by collecting terms involving ν. This is done element-wise as follows:

$$\begin{aligned}
\pi(\nu_i | z,\beta,h,\sigma,\gamma) &\propto \nu_i^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left[\frac{(z_i - x_i'\beta - A\nu_i - C|\gamma|h_i)^2}{\sigma B\nu_i}\right] - \frac{\nu_i}{\sigma}\right\} \\
&\propto \nu_i^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left[\frac{(z_i - x_i'\beta - C|\gamma|h_i)^2}{\sigma B}\nu_i^{-1} + \left(\frac{A^2}{\sigma B} + \frac{2}{\sigma}\right)\nu_i\right]\right\} \\
&\propto \nu_i^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left[a_i\nu_i^{-1} + b\nu_i\right]\right\},
\end{aligned}$$

which is recognized as the kernel of a GIG distribution where

$$a_i = \frac{(z_i - x_i'\beta - C|\gamma|h_i)^2}{\sigma B} \quad \text{and} \quad b = \left(\frac{A^2}{\sigma B} + \frac{2}{\sigma}\right).$$

Hence, $\nu_i | z, \beta, h, \sigma, \gamma \sim GIG(0.5,\ a_i,\ b)$ for $i = 1, 2, \ldots, n$.

(4) The conditional posterior of h is obtained element-wise from the complete posterior density (Eq. 15) conditional on γ and remaining parameters as follows:

$$\begin{aligned}
\pi(h_i \mid z, \beta, \nu, \sigma, \gamma) &\propto \exp\left\{ -\frac{1}{2}\left[\frac{(z_i - x_i'\beta - A\nu_i - C \mid \gamma \mid h_i)^2}{\sigma B \nu_i} + \frac{h_i^2}{\sigma^2} \right] \right\} \\
&\propto \exp\left\{ -\frac{1}{2}\left[\left(\frac{1}{\sigma^2} + \frac{C^2\gamma^2}{\sigma B \nu_i} \right) h_i^2 - \frac{2C \mid \gamma \mid (z_i - x_i'\beta - A\nu_i)}{\sigma B \nu_i} h_i \right] \right\} \\
&\propto \exp\left\{ -\frac{1}{2}\left[(\sigma_{h_i}^2)^{-1} h_i^2 - 2(\sigma_{h_i}^2)^{-1} \mu_{h_i} h_i \right] \right\} \\
&\propto \exp\left\{ -\frac{1}{2}(\sigma_{h_i}^2)^{-1}(h_i - \mu_{h_i})^2 \right\},
\end{aligned}$$

where the third line introduces the following notations:

$$(\sigma_{h_i}^2)^{-1} = \left(\frac{1}{\sigma^2} + \frac{C^2\gamma^2}{\sigma B \nu_i} \right) \quad \text{and} \quad \mu_{h_i} = \sigma_{h_i}^2 \left(\frac{C \mid \gamma \mid (z_i - x_i'\beta - A\nu_i)}{\sigma B \nu_i} \right),$$

and the fourth line adds and subtracts $(\sigma_{h_i}^2)^{-1}\mu_{h_i}^2$ to complete the square. The last expression is recognized as the kernel of a half-normal distribution, and hence, $h_i \mid z, \beta, \nu, \sigma, \gamma \sim N^+(\mu_{h_i}, \sigma_{h_i}^2)$ for $i = 1, 2, \ldots, n$.

(5) The transformed cut-point δ is sampled from the full likelihood (Eq. 9), marginally of (z, ν, h). The proposed values are generated from a random-walk chain, $\delta' = \delta_c + u$, where $u \sim N(0_{J-3}, \iota_2^2 \hat{D}_2)$, ι_2 is a tuning parameter, and $\hat{D}_2$ denotes negative inverse Hessian, obtained by maximizing the log-likelihood with respect to δ. Given the current value δ_c, the proposed value δ' is accepted with MH probability,

$$\alpha_{MH}(\delta_c, \delta') = \min\left\{ 0,\ \ln\left[\frac{f(y \mid \beta, \sigma, \gamma, \delta')\pi(\beta, \sigma, \gamma, \delta')}{f(y \mid \beta, \sigma, \gamma, \delta_c)\pi(\beta, \sigma, \gamma, \delta_c)} \right] \right\},$$

otherwise, the current value δ_c is repeated. The variance of u may be tuned as required for an appropriate step-size and acceptance rate.

(6) The full conditional density of the latent variable z is a truncated normal distribution where the cut-point vector ξ is obtained based on one-to-one mapping with δ. Hence, z is sampled as $z_i \mid \beta, \nu, h, \sigma, \gamma, \delta, y \sim TN_{(\xi_{j-1}, \xi_j)}(x_i'\beta + A\nu_i + C \mid \gamma \mid h_i, \sigma B \nu_i)$ for $i = 1, \ldots, n$ and $j = 1, \ldots, J$.