Flexible Bayesian Quantile Analysis of Residential Rental Rates

Ivan Jeliazkov
University of California, Irvine

Shubham Karnawat
University of California, Irvine

Mohammad Arshad Rahman
Indian Institute of Technology Kanpur

Angela Vossmeyer
Claremont McKenna College
and NBER*

August 21, 2025

Abstract

This article develops a random effects quantile regression model for panel data that allows for increased distributional flexibility, multivariate heterogeneity, and time-invariant covariates in situations where mean regression may be unsuitable. Our approach is Bayesian and builds upon the generalized asymmetric Laplace distribution to decouple the modeling of skewness from the quantile parameter. We derive two efficient simulation-based estimation algorithms, demonstrate their properties and performance in targeted simulation studies, and employ them in the computation of marginal likelihoods to enable formal Bayesian model comparisons. The methodology is applied in a study of U.S. residential rental rates following the Global Financial Crisis. Our empirical results provide interesting insights on the interaction between rents and economic, demographic and policy variables, weigh in on key modeling features, and overwhelmingly support the additional flexibility at all quantiles and across several sub-samples. The practical differences that arise as a result of allowing for flexible modeling can be nontrivial, especially for quantiles away from the median.

`Keywords:` Bayesian inference, generalized asymmetric Laplace distribution, Markov chain Monte Carlo, panel data, rental markets.

1 Introduction

This paper aims to provide complementary methodological and empirical contributions to the quantile regression literature. On the methodological side, we develop a flexible Bayesian approach to random effects quantile regression based on a generalization of the asymmetric Laplace (AL) distribution, specify an efficient Markov chain Monte Carlo (MCMC) estimation algorithm, and present methods for formal model comparison of various nested and non-nested models that also enable us

*Email addresses: `ivan@uci.edu`, `skarnawa@uci.edu`, `marshad@iitk.ac.in`, and `angela.vossmeyer@cmc.edu`.

to assess the importance of flexible modeling. Our methods are readily motivated by the econometric challenges of studying U.S. residential rental rates and their dependence on unemployment and mortgage policies following the Global Financial Crisis. Our investigation deals with key features of the data, including considerable zip-code-level heterogeneity and skewness in the distribution of rents. The separation of modeling features into those that are practically relevant from those that are not warranted in this context is handled by model comparison.

Koenker and Bassett (1978) introduced quantile regression as a minimization problem involving an asymmetrically weighted linear loss function, but subsequent work has noted the duality between that approach and modeling through a likelihood function built on the AL distribution (Koenker and Machado, 1999; Yu and Moyeed, 2001). The latter approach becomes very potent when the AL distribution is expressed as a mixture of normal and exponential distributions (Kozumi and Kobayashi, 2011). The mixture formulation permits estimation by simple, yet powerful, MCMC algorithms, and has enabled extensions of the quantile methodology to a variety of other settings including censored data (Kozumi and Kobayashi, 2011), binary data (Benoit and Poel, 2010; Ojha and Rahman, 2021), ordinal outcomes (Rahman, 2016; Alhamzawi, 2016; Maheshwari and Rahman, 2023), linear mixed models (Luo et al., 2012), and panels of binary (Rahman and Vossmeyer, 2019; Bresson et al., 2021), ordinal (Alhamzawi and Ali, 2018), or dynamic censored data (Kobayashi and Kozumi, 2012). Recent work by Goncalves et al. (2022) considered extensions to the case of dynamic quantile linear models. Geraci and Bottai (2007) propose a Monte Carlo EM algorithm for longitudinal data based on the AL likelihood and demonstrate efficiency gains relative to the regularized estimator of (Koenker, 2004), whereas Geraci and Bottai (2014) extend the approach to multiple random effects. Application of the AL distribution has unlocked a plethora of research opportunities, proving particularly valuable in longitudinal (panel) data settings, where the absence of appropriate data transformations to eliminate the latent heterogeneity has led to theoretical and computational challenges in frequentist estimation (see, e.g., Koenker, 2004, 2005; Abrevaya and Dahl, 2008; Lamarche, 2010; Galvao, 2011; Kato et al., 2012; Galvao et al., 2020).

In this paper, we extend current research in four directions. To improve the flexibility of the Bayesian quantile regression approach and dispense with certain limitations inherent in the AL specification, our modeling employs the generalized asymmetric Laplace (GAL) distribution of Yan and Kottas (2017). While the skewness of the AL distribution is completely determined once a quantile is chosen and the mode of the distribution is always fixed at the value of the location parameter, the GAL distribution circumvents these limitations by introducing a shape parameter into the mean of the normal kernel in the AL mixture representation. This allows for considerable flexibility with minimal tuning, identification or computational complexity, while also maintaining the AL distribution as an important special case. We further generalize our approach by allowing for heterogeneity in the skewness parameter across individuals. Heterogeneity can thus be captured in higher moments of the error distribution.

In adapting GAL modeling to the random effects panel setting, our second focus is on developing an efficient MCMC estimation algorithm that offers a variety of algorithmic improvements through suitable transformations of the mixture variables, block sampling of scale and shape parameters, and block sampling of the individual-specific and common effect parameters (cf. Nascimento and Goncalves, 2021). These aspects of the algorithm eliminate the problem of high autocorrelations in the MCMC draws, but in addition, they are applicable to the MCMC estimation of models based on the simpler AL distribution (cf. Luo et al., 2012) while also allowing for correlated random effects. Third, for both the GAL and AL panel data models, we aim to enable formal model comparison through marginal likelihoods, which, with few exceptions (e.g., Kobayashi and Kozumi, 2012; Maheshwari and Rahman, 2023) is broadly lacking in the quantile literature. We do so by adapting the methods of Chib (1995) and Chib and Jeliazkov (2001) to the building blocks of the proposed MCMC samplers. Several simulation studies carefully examine the properties and practical appeal of the proposed modeling, estimation and model comparison techniques.

Finally, the empirical contribution of the paper involves the study of U.S. residential rental rates during the recovery period following the Global Financial Crisis. We construct a novel data set that includes median rental rates in $14,533$ zip codes from 2010 to 2016, as well as zip-code-level economic, demographic, mortgage, and tax policy controls. Our methodology is particularly appealing in this setting because housing prices and rents are heavily skewed and heterogeneous across regions. For instance, from 2010 to 2016, the Cleveland MSA region's change in "All Transaction House Price Index" was about 8.42, whereas the San Francisco MSA region's change was about 156.[1] Skewness, along with heterogeneity in economic outcomes, has been identified as an important driver of public policy and political economy considerations (Benhabib and Bisin, 2018).

[1]Based on data from the FRED database at the Federal Reserve Bank of St. Louis.

The data reveal a positive impact of unemployment on residential rental rates as uncertain job prospects reduce the willingness and ability of households to commit to homeownership and instead shift demand towards rental units. We also find that home mortgage deductions decrease rental prices by making homeownership more attractive. This finding is particularly relevant in the context of the Tax Cuts and Jobs Act, which was passed in 2017. The law lowered mortgage deductions, suggesting that one consequence of the policy change is an expected rise in rents. Lastly, model comparisons across many quantiles and samples reveal that the data overwhelmingly support the more flexible GAL modeling framework, particularly when we allow for heterogeneity in the skewness parameter.

The remainder of the paper is organized as follows. In Section 2, we present the proposed modeling, estimation and model comparison framework. This section also presents improved algorithms for the simpler AL-based model. Section 3 illustrates the proposed algorithms in multiple simulation studies. Section 4 describes the data, presents our rental rates application and discusses the results, while Section 5 concludes.

2 Methodology

This section introduces our proposed model, discusses the challenges associated with its estimation, and presents an MCMC estimation algorithm for the GAL model that also leads on an improved algorithm for estimating simpler AL-based models. Further, we introduce a model extension that incorporates heterogeneity in the skewness parameter. Lastly, we discuss the problem of estimating marginal likelihoods for the purpose of enabling formal model comparisons.

2.1 The Flexible Random Effects Quantile (FREQ) Model

We focus on a panel data model which takes the form

$$y_{it} = x'_{it}\beta_{p_0} + z'_{it}\alpha_i + \varepsilon_{it}, \qquad i = 1, \ldots, n, \quad t = 1, \ldots, T_i, \tag{1}$$

where y_{it} denotes the t-th response on the i-th unit, x_{it} is a k vector of covariates, β_{p_0} is a k vector of common parameters at the p_0-th quantile (henceforth, simply β), z_{it} is an l vector of variables with $z_{it} \subseteq x_{it}$ and a corresponding l vector of subject-specific random effects α_i (with $E(\alpha_i) = 0$

for identification purposes), that induces dependence among the observations on unit i.[2]

Although not immediately obvious from the notation, the setup is rather general and can capture dynamics, unknown covariate functions, and correlated random effects, depending on what is included in x_{it} (and potentially also in $z_{it} \subseteq x_{it}$). In particular, dynamic modeling can be pursued by including lags of y_{it} in x_{it} and ensuring that the lag coefficients satisfy stationarity. Flexible functional modeling for some covariate s can be implemented through a set of basis functions $\mathcal{B} = \{b_1, \ldots, b_m\}$, e.g., B-splines, natural splines, truncated power series, wavelets, etc., (Ruppert et al., 2003) so that $f(s) = \sum_{j=1}^{m} b_j(s)\delta_j$, in which case x_{it} includes $(b_1(s_i), \ldots, b_m(s_i))'$, while $(\delta_1, \ldots, \delta_m)'$ becomes part of the regression parameter vector β. In addition, correlated random effect models where the heterogeneity can be correlated with certain observed covariates is handled by interacting those covariates with z_{it} and including the result in x_{it} (see, e.g., Chamberlain, 1984; Mundlak, 1978; Chib and Jeliazkov, 2006). Random effect models are also indispensable because of their versatility – they can address both univariate and multivariate heterogeneity, can be used with time-invariant covariates, and are broadly applicable in models for different types of outcomes. In contrast, data transformations such as mean- or first-differencing, which underlie conventional "fixed effects" estimators in econometrics remove heterogeneity in the intercept only but do not handle variation in slopes, wipe out any time-invariant covariates, and are inapplicable in non-linear models. Critically, however, such transformations can not be employed in our context since differencing a quantile equation in levels does not lead to an analogous quantile regression on the differences; pursuing the alternative of estimating all individual effects has led to theoretical and computational challenges in frequentist estimation (Koenker, 2004; Kato et al., 2012; Galvao et al., 2020).

We parameterize the model in Equation (1) by letting $\varepsilon_{it} \overset{iid}{\sim} \mathrm{GAL}(0, \sigma, p_0, \gamma)$, using the quantile-fixed GAL distribution (Yan and Kottas, 2017)—a generalization stemming from the mixture representation of the AL distribution—to decouple the modeling of skewness from the quantile parameter. A variable s is said to follow a quantile-fixed GAL distribution, i.e., $s \sim \mathrm{GAL}(\mu, \sigma, p_0, \gamma)$, where μ, σ, p_0, and γ represent the location, scale, quantile, and skewness parameters, respectively,

[2] An unfortunate rift in terminology has persisted between statistics and econometrics in the panel (longitudinal) context. In statistics, β and $\{\alpha_i\}$ are called fixed and random effects, respectively, because the former do not vary with i, whereas the latter are subject-specific. In econometrics, these terms are used to distinguish between alternative ways of dealing with $\{\alpha_i\}$ – fixed effects estimators remove the heterogeneity (if possible) by data transformations such as mean- or first-differencing, whereas random effects estimators model the $\{\alpha_i\}$ explicitly through a distribution.

if it has density

$$
\begin{aligned}
f_{GAL}(s|\mu,\sigma,p_0,\gamma) = \frac{2p(1-p)}{\sigma} \Bigg(\Bigg[& \Phi\left(-s^* \frac{p_{\gamma+}}{|\gamma|} + \frac{p_{\gamma-}}{p_{\gamma+}}|\gamma|\right) - \Phi\left(\frac{p_{\gamma-}}{p_{\gamma+}}|\gamma|\right) \Bigg] \\
& \times \exp\left\{ - s^* p_{\gamma-} + \frac{\gamma^2}{2}\left(\frac{p_{\gamma-}}{p_{\gamma+}}\right)^2 \right\} I\left(\frac{s^*}{\gamma} > 0\right) \\
& + \Phi\left(-|\gamma| + s^* \frac{p_{\gamma+}}{|\gamma|} I\left(\frac{s^*}{\gamma} > 0\right)\right) \exp\left\{ - s^* p_{\gamma+} + \frac{\gamma^2}{2} \right\} \Bigg),
\end{aligned}
\tag{2}
$$

where $s^* = \frac{s-\mu}{\sigma}$, $p \equiv p(\gamma, p_0) = I(\gamma < 0) + [p_0 - I(\gamma < 0)]/g(\gamma)$, $p_{\gamma+} = p - I(\gamma > 0)$ and $p_{\gamma-} = p - I(\gamma < 0)$, $g(\gamma) = 2\Phi(-|\gamma|)\exp(\gamma^2/2)$ 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 $g(\gamma) = p_0$; additional discussion is offered in Rahman and Karnawat (2019, Section 2). The term "quantile-fixed" refers to the fact that the density in Equation (2) satisfies $\int_{-\infty}^{\mu} f_{GAL}(\varepsilon_{it}|\mu,\sigma,p_0,\gamma)d\varepsilon_{it} = p_0$. The special case of AL density results when $\gamma = 0$, which is seen more clearly from the mixture representation discussed shortly.

Figure 1 offers a visualization of the differences between the quantile-fixed GAL and AL densities. The figure shows three different quantiles when $\sigma = 1$, the standard case. 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. Additionally, at the median $p_0 = 0.50$, the GAL distribution can be positively ($\gamma < 0$) or negatively ($\gamma > 0$) skewed and can have tails that are heavier or thinner than the AL distribution. These characteristics make GAL significantly more flexible than AL, but the value of the extra adaptability will generally be an application-specific empirical question.

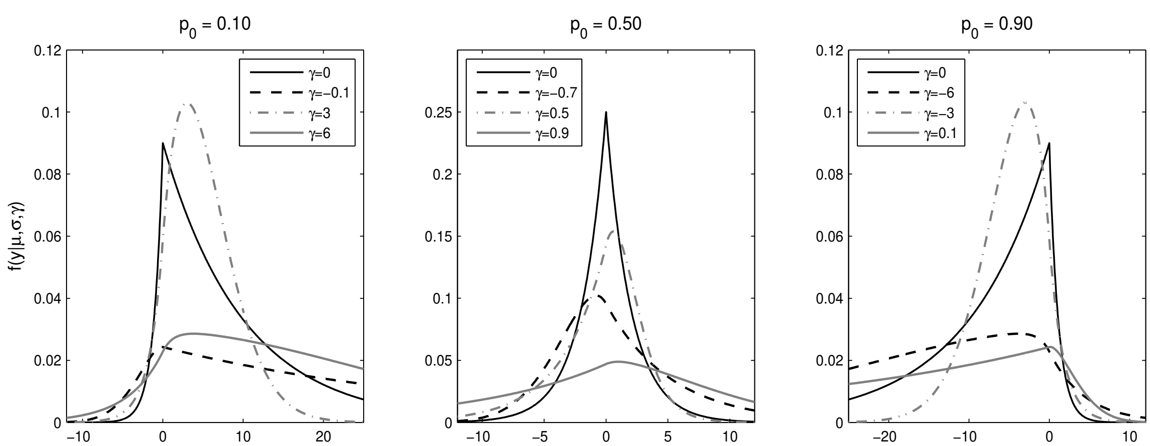

Figure 1: Probability density plots of the AL ($\gamma = 0$) and GAL ($\gamma \neq 0$) distributions.

Because of the additional flexibility that the GAL distribution offers over the AL distribution, we refer to the model based on the former as the Flexible Random Effects Quantile (FREQ) model, and the model based on the latter as the Random Effects Quantile (REQ) model.

The distributional assumption on the error term, $\varepsilon_{it} \overset{iid}{\sim} \text{GAL}(0, \sigma, p_0, \gamma)$ implies that $y_{it}|\alpha_i \overset{ind}{\sim} \text{GAL}(x'_{it}\beta + z'_{it}\alpha_i, \sigma, p_0, \gamma)$ for $i = 1, \ldots, n$ and $t = 1, \ldots, T_i$. Assuming a density $f(\alpha|\Omega)$ for the random effects, the complete data likelihood can be expressed as

$$f(y, \alpha|\beta, \sigma, \gamma, \Omega) = \prod_{i=1}^{n} \left[\left\{ \prod_{t=1}^{T_i} f_{GAL} \left(y_{it}|x'_{it}\beta + z'_{it}\alpha_i, \sigma, p_0, \gamma \right) \right\} f(\alpha_i|\Omega) \right], \tag{3}$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$, $y = (y_1, \ldots, y_n)$ with each $y_i = (y_{i1}, \ldots, y_{iT_i})'$ for $i = 1, \ldots, n$. The density $f(\alpha_i|\Omega)$ can be any suitable distribution, but is typically assumed normal (Luo et al., 2012), e.g., here we let $\alpha_i|\Omega \overset{iid}{\sim} N(0_l, \Omega)$ for $i = 1, \ldots, n$. The complete data likelihood in Equation (3) can be combined with priors on the parameters to obtain the joint posterior distribution, but this posterior does not yield conditional densities suitable for MCMC sampling. Hence, we utilize the GAL mixture representation of Yan and Kottas (2017).

For $\varepsilon_{it} \sim \text{GAL}(0, \sigma, p_0, \gamma)$, the mixture representation can be expressed as,

$$\varepsilon_{it} = \sigma A \omega_{it} + \sigma C|\gamma| s_{it} + \sigma (B\omega_{it})^{\frac{1}{2}} u_{it}, \tag{4}$$

where $s_{it} \sim N^+(0, 1)$, $\omega_{it} \sim \mathcal{E}(1)$, $u_{it} \sim N(0, 1)$, $A \equiv A(p) = \frac{1-2p}{p(1-p)}$, $B \equiv B(p) = \frac{2}{p(1-p)}$, $C = [I(\gamma > 0) - p]^{-1}$, and p is as defined earlier. Here, $N^+, \mathcal{E}, N$ denote the half-normal, exponential, and normal distributions, respectively. Note that the GAL mixture representation is obtained by introducing a shape parameter into the mean of the normal kernel in the normal-exponential mixture of the AL distribution and mixing with respect to a half-normal distribution; consequently, it reduces to an AL mixture distribution when $\gamma = 0$. Substituting the representation (4) into Equation (1), employing the transformation $h_{it} = \sigma s_{it}$, $\nu_{it} = \sigma \omega_{it}$ (Rahman and Karnawat, 2019), and stacking the observations for each individual, we have

$$y_i = X_i\beta + Z_i\alpha_i + A\nu_i + C|\gamma|h_i + \Lambda_i^{1/2} u_i, \qquad i = 1, \ldots, n, \tag{5}$$

where $y_i = (y_{i1}, \cdots, y_{iT_i})'$, $X_i = (x'_{i1}, \cdots, x'_{iT_i})'$, $Z_i = (z'_{i1}, \cdots, z'_{iT_i})'$, $\nu_i = (\nu_{i1}, \cdots, \nu_{iT_i})'$, $h_i = (h_{i1}, \cdots, h_{iT_i})'$, $\Lambda_i = \text{diag}(\sigma B \nu_i)$ and $u_i = (u_{i1}, \cdots, u_{iT_i})'$. Letting $\Theta = (\beta, \alpha, \nu, h, \sigma, \gamma, \Omega)$, we therefore have that $f(y_i|\Theta) = \prod_{i=1}^{n} f_N \left(y_i|X_i\beta + Z_i\alpha_i + A\nu_i + C|\gamma|h_i, \Lambda_i \right)$, which is combined with the priors

$$\beta \sim N(\beta_0, B_0), \qquad \sigma \sim IG\left(n_0/2, d_0/2\right), \qquad \gamma \sim Unif(L, U), \qquad \Omega \sim IW(r_0, R_0) \tag{6}$$

to obtain the joint posterior distribution.[3] Based on this modeling framework, we develop an estimation algorithm for the FREQ model in Section 2.2 and consider computation of its marginal likelihood in Section 2.4.

Before delving into these details, we emphasize key features of the methodology, suggest avenues for potential extensions and further research, and consider several caveats. One observation is that the GAL mixture formulation facilitates model extensions to censored, binary, and ordinal data settings, building upon the literature referenced in Section 1, with estimation involving the integration of appropriate data augmentation steps into the algorithms presented in Section 2.2. Another point is that by examining a range of quantiles in the panel context, the methodology can capture dependencies across the entire conditional distribution of y, and in particular can provide insights into the variability of outcomes within and across clusters and the behavior of heterogeneity across quantiles. This can be contrasted with frequentist approaches building upon Koenker (2004), where the heterogeneity is not quantile-dependent, unlike the remaining model parameters, which vary over quantiles (Lamarche, 2021, offers a recent review).

Despite the benefits of accounting for heterogeneity in the panel context, its presence poses challenges for a number of possible extensions. For instance, there is a growing Bayesian literature on the simultaneous estimation of different quantiles (see, e.g., Taddy and Kottas, 2010; Yang and Tokdar, 2017; Das and Ghosal, 2017; Chen and Tokdar, 2021), but adapting these methods to the case with multiple heterogeneous effects across individuals and quantiles greatly increases the dimensionality of the quantile surfaces and raises both computational and modeling challenges that require further exploration. Extensions are also possible based on Bayesian approaches for flexible distributional modeling, including scale, finite, and infinite mixture models that can be employed in modeling the random effects distribution (see, e.g., Ferguson, 1973; Antoniak, 1974; Andrews and Mallows, 1974; Escobar and West, 1995; Neal, 2000; Koop and Tobias, 2004; Ho and Hu, 2008; Dunson, 2009; Kyung et al., 2010, among others). However, owing to identification issues and the curse of dimensionality, caution is necessary in short panels (small T), when the dimension of the random effects l is large, or with a small number of clusters n. In these instances, the individual

[3]The inverse gamma (IG) and inverse Wishart (IW) distributions are parameterized as in Gelman et al. (2013).

effects and their distribution may be poorly identified, and the results may exhibit sensitivity or require delicate choices of tuning parameters. Such issues may be further exacerbated when the modeling hierarchy is extended, e.g., in models for discrete data.

The approach offered in this paper is also adaptable to spatial panel data models. Recent work on spatial modeling ...

Work has also be done on flexible distributional modeling in both median and arbitrary quantile regression (Kottas and Gelfand, 2001; Hanson and Johnson, 2002; Kottas and Krnjajić, 2009; Reich et al., 2010). As discussed earlier, the GAL distribution employed here is an example of a mixture distribution that offers considerable parametric flexibility over traditional AL models at the cost of a single additional parameter, which aids in interpretability and identification. Unlike AL mixtures, GAL does not constrain the mode or skewness once a quantile is chosen. Instead, it strikes a balance by offering flexibility while limiting computational complexity, identification, tuning, and interpretation concerns. When analyzing each quantile, the more general modeling of tail behavior and skewness allows for additional flexibility, greater relevance, and improved fit in settings where data are skewed, exhibit power laws, or present asymmetries. Examples include income, bank assets, social networks, home prices, and rental rates, the latter being the focus of our application. Lastly, because of the parametric flexibility, we are able to extend the model by allowing for heterogeneity in the skewness parameter. This extension is discussed in Section 2.3.

2.2 Estimation

Our proposed MCMC sampler for the FREQ model is presented in Algorithm 1, and the relevant derivations are offered in Appendix A. Part of the challenge in designing an efficient algorithm for our implementation is recognizing the problem-specific dependencies among parameters. Significant improvements in the performance of the Markov chain can be afforded by careful blocking and joint simulation of parameters that are correlated in the posterior. The strategy has been employed with different types of samplers in a variety of hierarchical, multimodal, or high-dimensional settings to improve mixing and reduce the inefficiency factors of the Markov chain (Roberts and Sahu, 1997; Chib and Carlin, 1999; Chib and Ramamurthy, 2010; Greenberg, 2012; Turek et al., 2017). Successful simulation designs improve the quality of the MCMC sample and reduce inefficiency factors by more than any increase in computational costs associated with blocking, thereby leading to more efficient sampling of the entire model.

One instance where block sampling leads to marked improvements in MCMC performance in our implementation involves the sampling of α and β. In particular, because $z_{it} \subseteq x_{it}$ and the two sets of parameters enter additively in Equation (3), they will be correlated in the posterior by construction. Therefore, in Step 1 of Algorithm 1, we proceed by joint simulation of (β, α) in a single block by first sampling β from a conditional posterior marginally of α, and then drawing α conditionally on β and everything else (Chib and Carlin, 1999).

Another instance where blocking is helpful includes the simulation of the scale and shape parameters σ and γ, respectively. Although drawing each of them separately may seem desirable because of conjugacy or dimensionality considerations, it can be seen from Equation (2) that these parameters are closely related in the GAL density. We therefore proceed by simulating (σ, γ) jointly and marginally of (ν, h) in Step 2 of Algorithm 1 using a random-walk Metropolis-Hastings (MH) step (Tierney, 1994; Chib and Greenberg, 1995) from a target density proportional to the product of Equations (3) and (6), with proposal values coming from a bivariate truncated normal distribution. Joint sampling of (β, α) in Step 1 and (σ, γ) in Step 2 comes at a nominal computational cost but is critical in reducing the MCMC autocorrelations and improving the efficiency of the algorithm.

The remaining steps of the estimation algorithm are more straightforward. In Step 3, we draw the mixture variable ν from a generalized inverse Gaussian (GIG) distribution using the technique in Devroye (2014). In Step 4 of the algorithm, the mixture variable h is sampled from a half-normal distribution with updated hyperparameters. Finally, because of conditional conjugacy, in Step 3 of Algorithm 1 the covariance matrix of the random effects Ω is sampled from an inverse-Wishart distribution with updated hyperparameters. Appendix A contains further derivations and computational details on the Steps of Algorithm 1.

To study the practical utility of the FREQ model, we also estimate the simpler and more established REQ model using Algorithm 2. This sampling algorithm has two important improvements over the sampler proposed in Luo et al. (2012). First, (β, α) are sampled in a single block which significantly lowers the MCMC autocorrelations and improves mixing. Because it can achieve lower inefficiency factors, the number of MCMC draws in Algorithm 2 can be reduced, thereby decreasing computational burdens and run times. Second, we correct the updating for σ by including the terms involving the exponential variable in the posterior updating of the hyperparameters (cf.

Algorithm 1 (Sampling in the FREQ model)

---

(1) Sample $\beta, \alpha | y, \nu, h, \sigma, \gamma, \Omega$ in a single block as follows:

(a) Sample $\beta | y, \nu, h, \sigma, \gamma, \Omega \sim N(\tilde{\beta}, \tilde{B})$ marginally of α, where, given $V_i = Z_i \Omega Z_i' + \Lambda_i$,

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

(b) Sample $\alpha_i | y, \beta, \nu, h, \sigma, \gamma, \Omega \sim N(\tilde{\alpha}_i, \tilde{A}_i)$, where

$$\tilde{A}_i = \left(Z_i' \Lambda_i^{-1} Z_i + \Omega^{-1} \right)^{-1} \quad \text{and} \quad \tilde{\alpha}_i = \tilde{A}_i \left(Z_i' \Lambda_i^{-1} (y_i - X_i \beta - A\nu_i - C|\gamma| h_i) \right).$$

(2) Jointly sample $\sigma, \gamma | y, \beta, \alpha, \sigma, \gamma, \Omega$ marginally of (ν, h) using a Metropolis-Hasings step. Given the current values (σ_c, γ_c), generate a proposed draw $(\sigma^\dagger, \gamma^\dagger) \sim TN_{(0,\infty)\times(L,U)}\big((\sigma_c, \gamma_c), \iota^2 \hat{D}\big)$ where ι is a tuning factor and $\hat{D}$ is obtained by tailoring an approximation to the target at the start of the Markov chain. Accept $(\sigma^\dagger, \gamma^\dagger)$ as the next MCMC draw with probability

$$\alpha_{MH}(\sigma_c, \gamma_c; \sigma^\dagger, \gamma^\dagger | y, \cdot) = \min \left\{ 1, \left[\frac{f(y, \alpha | \beta, \sigma^\dagger, \gamma^\dagger)\, \pi(\beta, \sigma^\dagger, \gamma^\dagger)}{f(y, \alpha | \beta, \sigma_c, \gamma_c)\, \pi(\beta, \alpha, \sigma_c, \gamma_c)} \; \frac{q(\sigma_c, \gamma_c | (\sigma^\dagger, \gamma^\dagger), \iota^2 \hat{D})}{q(\sigma^\dagger, \gamma^\dagger | (\sigma_c, \gamma_c), \iota^2 \hat{D})} \right] \right\},$$

where $f(\cdot)$ is the complete data likelihood, $\pi(\cdot)$ denotes the prior density, and $q(\cdot)$ represents the proposal density. If $(\sigma^\dagger, \gamma^\dagger)$ is not accepted, (σ_c, γ_c) is repeated as the next MCMC draw.

(3) Sample $\nu_{it} | y_{it}, \beta, \alpha_i, h_{it}, \sigma, \gamma \sim GIG(\frac{1}{2}, \chi, \psi_{\nu_{it}})$ for all values of i and t, where

$$\chi = \left(\frac{A^2}{\sigma B} + \frac{2}{\sigma} \right) \quad \text{and} \quad \psi_{\nu_{it}} = \frac{(y_{it} - x_{it}'\beta - z_{it}'\alpha_i - C|\gamma| h_{it})^2}{\sigma B}.$$

(4) Sample $h_{it} | y_{it}, \beta, \nu_{it}, \sigma, \gamma \sim N^+(\mu_{h_{it}}, \sigma^2_{h_{it}})$ for all values of i and t, where,

$$\sigma^2_{h_{it}} = \left(\frac{1}{\sigma^2} + \frac{C^2 \gamma^2}{\sigma B \nu_{it}} \right)^{-1} \quad \text{and} \quad \mu_{h_{it}} = \sigma^2_{h_{it}} \left(\frac{C|\gamma| (y_{it} - x_{it}'\beta - z_{it}'\alpha_i - A\nu_{it})}{\sigma B \nu_{it}} \right).$$

(5) Sample $\Omega | \alpha \sim IW(\tilde{\omega}, \tilde{R})$, where $\tilde{\omega} = n + r_0$ and $\tilde{R} = \sum_{i=1}^{n} \alpha_i \alpha_i' + R_0$.

Kozumi and Kobayashi, 2011). The resulting MCMC algorithm is fast, efficient, and maintains the tractability of the sampling distributions.

In closing, we note that the algorithms considered here were developed for the general model with multivariate heterogeneity present in both the intercept and some subset of the slope parameters,

Algorithm 2 (Sampling in the REQ model)

---

(1) Sample $\beta, \alpha | y, \nu, \sigma, \Omega$ in a single block:

(a) Sample $\beta | y, \nu, \sigma, \Omega \sim N(\tilde{\beta}, \tilde{B})$, marginally of α, where

$$\tilde{B} = \left(\sum_{i=1}^{n} X_i' V_i^{-1} X_i + B_0^{-1} \right)^{-1} \quad \text{and} \quad \tilde{\beta} = \tilde{B} \left(\sum_{i=1}^{n} X_i' V_i^{-1} (y_i - A\nu_i) + B_0^{-1} \beta_0 \right).$$

(b) Sample $\alpha_i | y, \beta, \nu, \sigma, \Omega \sim N(\tilde{\alpha}_i, \tilde{A}_i)$, where

$$\tilde{A}_i = \left(Z_i' \Lambda_i^{-1} Z_i + \Omega^{-1} \right)^{-1} \quad \text{and} \quad \tilde{a}_i = \tilde{A}_i \left(Z_i' \Lambda_i^{-1} (y_i - X_i \beta - A\nu_i) \right).$$

(2) Sample $\nu_{it} | y_{it}, \beta, \alpha_i, \sigma, \gamma \sim GIG(\frac{1}{2}, \chi, \psi_{\nu_{it}})$, with

$$\chi = \left(\frac{A^2}{\sigma B} + \frac{2}{\sigma} \right) \quad \text{and} \quad \psi_{\nu_{it}} = \frac{(y_{it} - x_{it}'\beta - z_{it}'\alpha_i)}{\sigma B}.$$

(3) Sample the scale parameter $\sigma | y, \beta, \alpha, \nu \sim IG\left(\tilde{n}/2, \tilde{d}/2\right)$, where,

$$\tilde{n} = 3 \sum_{i=1}^{n} T_i + n_0 \qquad \text{and} \qquad \tilde{d} = \sum_{i=1}^{n} \sum_{t=1}^{T_i} \left\{ \frac{(y_{it} - x_{it}'\beta - z_{it}'\alpha_i - A\nu_{it})^2}{B\nu_{it}} + 2\nu_{it} \right\} + d_0.$$

(4) Sample $\Omega | \alpha \sim IW(\tilde{\omega}, \tilde{R})$, where, $\tilde{\omega} = n + r_0$ and $\tilde{R} = \sum_{i=1}^{n} (\alpha_i \alpha_i') + R_0$.

---

but they trivially simplify to the case of heterogeneity in the intercept only. For a scalar Ω, the inverse Wishart density is equivalent to an inverse gamma; therefore, we can continue to work with the prior on Ω in Equation (6) and the full-conditionals in Step 5 of Algorithm 1 and Step 4 of Algorithm 2, or we can specify an equivalent inverse gamma prior $\Omega \sim IG(c_1/2, d_1/2)$ and replace the full conditionals in Algorithms 1 and 2 with $\Omega | \alpha \sim IG(\tilde{c}_1/2, \tilde{d}_1/2)$, where $\tilde{c}_1 = n + c_1$ and $\tilde{d}_1 = \sum_{i=1}^{n} \alpha_i^2 + d_1$.[4] No conceptual changes are necessary elsewhere in these algorithms.

[4]Inverse gamma priors, for example, are specified in our study of intercept heterogeneity in the rental rates application in Section 4.

2.3 Generalized FREQ: Heterogeneity in the Skewness Parameter

In this section, we propose a generalization of the FREQ model (1), referred to as the GFREQ model, by introducing cluster-specific heterogeneity in the skewness parameter, denoted by γ_i. By allowing the skewness parameter to vary with the cluster index i, the model accommodates variability in the shape of the error distribution, thereby enhancing its flexibility. This approach to modeling cluster-specific skewness represents a novel contribution and has not been explored in the existing literature, including Yan and Kottas (2017) and Rahman and Karnawat (2019).

The GFREQ model, for unit i at time t, can be written as follows:

$$y_{it} = x'_{it}\beta_{p0} + z'_{it}\alpha_i + \varepsilon_{it}, \qquad \varepsilon_{it} \overset{iid}{\sim} \text{GAL}(0, \sigma, p_0, \gamma_i), \tag{7}$$

and when stacked over time for each unit i, yields the following formulation,

$$y_i = X_i\beta + Z_i\alpha_i + A_i\nu_i + C_i|\gamma_i|h_i + \Lambda_i^{1/2}u_i, \qquad i = 1, \ldots, n, \tag{8}$$

where $A \equiv A_i(p_i) = \frac{1-2p_i}{p_i(1-p_i)}$, $B \equiv B_i(p_i) = \frac{2}{p_i(1-p_i)}$, $C_i = [I(\gamma > 0) - p_i]^{-1}$, $p_i \equiv p(\gamma_i, p_0) = I(\gamma_i < 0) + [p_0 - I(\gamma_i < 0)]/g(\gamma_i)$ and $g(\gamma_i) = 2\Phi(-|\gamma_i|)\exp(\gamma_i^2/2)$.

Thereafter, the likelihood is $f(y|\Theta) = \prod_{i=1}^{n} f_N\left(y_i|X_i\beta + Z_i\alpha_i + A_i\nu_i + C_i|\gamma_i|h_i, \Lambda_i\right)$, can be combined with the following priors: $\beta \sim N(\beta_0, B_0)$, $\sigma \sim IG\left(n_0/2, d_0/2\right)$, $\gamma \sim Unif(L, U)$, $\Omega \sim IW(r_0, R_0)$ to arrive at joint posterior distribution. To accommodate this additional form of heterogeneity, Algorithm 1 can be adapted by simulating (σ, γ) jointly in a single block and marginally of (ν, h) in Step 2 of the algorithm. This is another instance where blocking is crucial. Although drawing σ and $\{\gamma_i\}$ separately may seem desirable because of conjugacy or dimensionality considerations, it can be seen from Equation (2) that these parameters are closely related in the GAL density.

Algorithm 3 details the full sampler. Notably, we simulate σ using a random-walk Metropolis-Hastings (MH) step from a target density proportional to the product of Equations (9) and (6), with proposal values coming from truncated normal distribution (Tierney, 1994; Chib and Greenberg, 1995),

$$f(y, \alpha|\beta, \sigma, \Omega) = \prod_{i=1}^{n}\left[\int_L^U f\left(y_i|X_i\beta + Z_i\alpha_i, \sigma, p_0, \gamma\right) f(\alpha_i|\Omega)\pi(\gamma_i)d\gamma_i\right]. \tag{9}$$

It then follows that we simulate $\{\gamma_i\}$ for each cluster unit i conditional on (β, α, σ and Ω) using random-walk MH, using the target density and truncated normal as before. It is important to note

Algorithm 3 (Sampling in the Generalized FREQ model)

---

(1) Sample $\beta, \alpha | y, \nu, s, \sigma, \gamma, \Omega$ in a single block as follows:

(a) Sample $\beta | y, \nu, s, \sigma, \gamma, \Omega \sim N(\tilde{\beta}, \tilde{B})$ marginally of α, where, given $V_i = Z_i \Omega Z_i' + \Lambda_i$,

$$\tilde{B} = \left(\sum_{i=1}^{n} X_i' V_i^{-1} X_i + B_0^{-1} \right)^{-1} \quad \text{and} \quad \tilde{\beta} = \tilde{B} \left(\sum_{i=1}^{n} X_i' V_i^{-1} (y_i - A_i \nu_i - C_i |\gamma_i| h_i) + B_0^{-1} \beta_0 \right).$$

(b) Sample $\alpha_i | y, \beta, \nu, s, \sigma, \gamma, \Omega \sim N(\tilde{\alpha}_i, \tilde{A}_i)$, where

$$\tilde{A}_i = \left(Z_i' \Lambda_i^{-1} Z_i + \Omega^{-1} \right)^{-1} \quad \text{and} \quad \tilde{\alpha}_i = \tilde{A}_i \left(Z_i' \Lambda_i^{-1} (y_i - X_i \beta - A_i \nu_i - C_i |\gamma_i| h_i) \right).$$

(2) Sample $\sigma, \gamma | y, \beta, \alpha$, in a single block as follows:

(a) Sample σ marginally of $\{\gamma_i\}_{i=1}^{n}$ using a Metropolis-Hastings step. Given current value σ_c, generate a proposed draw $\sigma^{\dagger} \sim TN_{(0,\infty)} \left(\sigma_c, \iota_\sigma^2 \hat{D}_\sigma \right)$ where ι is the tuning parameter and $\hat{D}_\sigma$ is obtained by tailoring an approximation to the target at the start of Markov chain. Accept $\sigma^{\dagger}$ with probability

$$\alpha_{MH}(\sigma_c; \sigma^{\dagger} | y, \cdot) = \min \left\{ 1, \left[\frac{\prod_{i=1}^{n} \int_L^U f(y_i, \alpha_i | \beta, \sigma^{\dagger}, \gamma_i) d\gamma_i}{\prod_{i=1}^{n} \int_L^U f(y_i, \alpha_i | \beta, \sigma, \gamma_i) d\gamma_i} \, \frac{\pi(\sigma^{\dagger})}{\pi(\sigma_c)} \, \frac{q(\sigma^c | \sigma^{\dagger}, \iota_\sigma^2 \hat{D}_\sigma)}{q(\sigma^{\dagger} | \sigma_c, \iota_\sigma^2 \hat{D}_\sigma)} \right] \right\},$$

where $f(y_i, \alpha_i | .)$ is the integrated GAL density, $\pi(\cdot)$ denotes the prior density, and $q(\cdot)$ represents the proposal density. If $(\sigma^{\dagger})$ is not accepted, (σ^c) is repeated.

(b) For each cluster i, sample $\gamma_i | y, \beta, \alpha_i, \sigma$ using a Metropolis-Hastings step. Generate a proposed draw $\gamma_i^{\dagger} \sim TN_{(L,U)}(\gamma_i^c, \iota_\gamma^2 \hat{D}_\gamma)$. Where ι_γ is a tuning factor and $\hat{\gamma}$ is obtained by tailoring an approximation to the target at the start of the Markov chain. Accept $\gamma_i^{\dagger}$ as the next MCMC draw with probability

$$\alpha_{MH}(\gamma_c; \sigma^{\dagger} | y, \cdot) = \min \left\{ 1, \left[\frac{f(y_i, \alpha_i | \beta, \sigma, \gamma_i^{\dagger}) \, \pi(\gamma_i^{\dagger})}{f(y_i, \alpha_i | \beta, \sigma, \gamma_i^{c}) \, \pi(\gamma_i^{\dagger})} \, \frac{q(\gamma_i^c | \gamma_i^{\dagger}, \iota^2 \hat{D})}{q(\gamma_i^{\dagger} | \gamma_i^c, \iota^2 \hat{D})} \right] \right\},$$

where $f(y_i, \alpha_i | .)$ is the GAL density given by (2), $\pi(\cdot)$ denotes the prior density, and $q(\cdot)$ represents the proposal density. If $(\gamma_i^{\dagger})$ is not accepted, (γ_i^c) is repeated.

(3) Sample $\nu_{it} | y_{it}, z_{it}, \beta, \alpha_i, s_{it}, \sigma, \gamma_i \sim GIG(\frac{1}{2}, \chi, \psi_{\nu_{it}})$ for all values of i and t, where

$$\chi = \left(\frac{A_i^2}{\sigma B_i} + \frac{2}{\sigma} \right) \quad \text{and} \quad \psi_{\nu_{it}} = \frac{(y_{it} - x_{it}' \beta - z_{it}' \alpha_i - C_i |\gamma_i| h_{it})^2}{\sigma B_i}.$$

(4) Sample $h_{it} | y_{it}, \beta, \nu_{it}, \sigma, \gamma_i \sim N^{+}(\mu_{h_{it}}, \sigma_{h_{it}}^2)$ for all values of i and t, where,

$$\sigma_{h_{it}}^2 = \left(\frac{1}{\sigma^2} + \frac{C_i^2 \gamma_i^2}{\sigma B_i \nu_{it}} \right)^{-1} \quad \text{and} \quad \mu_{s_{it}} = \sigma_{s_{it}}^2 \left(\frac{C_i |\gamma_i| (y_{it} - x_{it}' \beta - z_{it}' \alpha_i - A_i \nu_{it})}{\sigma B_i \nu_{it}} \right).$$

(5) Sample $\Omega | \alpha \sim IW(\tilde{\omega}, \tilde{R})$, where $\tilde{\omega} = n + r_0$ and $\tilde{R} = \sum_{i=1}^{n} \alpha_i \alpha_i' + R_0$.

---

that, both the integration task in Equation (9) and the sampling of γ_i can be efficiently parallelized across multiple machines, as the likelihood contributions of each cluster are independent and do not require sequential computation. This significantly reduces the computational burden associated with allowing heterogeneity in γ.

2.4 Bayesian Model Comparison and Marginal Likelihood Estimation

To properly address model uncertainty, Bayesian model comparison proceeds by representing the posterior model probability of model $\mathcal{M}_s$ given the data y as

$$\Pr(\mathcal{M}_s|y) \propto \Pr(\mathcal{M}_s)m(y|\mathcal{M}_s),$$

where $\Pr(\mathcal{M}_s)$ is the prior model probability and $m(y|\mathcal{M}_s)$ is the marginal likelihood of model $\mathcal{M}_s$, defined as the integral of the likelihood $f(y|\mathcal{M}_s, \Theta_s)$ with respect to the prior $\pi(\Theta_s|\mathcal{M}_s)$ over the parameters Θ_s of model $\mathcal{M}_s$, i.e., $m(y|\mathcal{M}_s) = \int f(y|\mathcal{M}_s, \Theta_s)\pi(\Theta_s|\mathcal{M}_s)\, d\Theta_s$. Chib (1995) notes that the marginal likelihood can also be expressed by Bayes' theorem as

$$m(y|\mathcal{M}_s) = \frac{f(y|\mathcal{M}_s, \Theta_s)\, \pi(\Theta_s|\mathcal{M}_s)}{\pi(\Theta_s|y, \mathcal{M}_s)}, \tag{10}$$

which is an identity that holds for any Θ_s in the parameter space. The numerator quantities in Equation (10) are typically available directly, and therefore the problem of marginal likelihood estimation is reduced to finding an estimate of the posterior ordinate in the denominator of Equation (10). To reduce variability, the posterior ordinate estimate is usually computed at a high-density point Θ_s^* (e.g., the posterior mean or mode).

The Bayesian model comparison methodology based on marginal likelihoods and their ratios, or Bayes factors, has well-known desirable properties – it leads to finite-sample model probabilities, does not require competing models to be nested, and has appealing asymptotic behavior that gives rise to the information criterion of Schwarz (1978). Another important, yet underappreciated, point is that marginal likelihoods provide a measure of sequential out-of-sample predictive fit, which can be seen by writing

$$\begin{aligned} m(y|\mathcal{M}_s) &= \prod_{i=1}^{n} m(y_i|\{y_j\}_{j<i}, \mathcal{M}_s) \\ &= \prod_{i=1}^{n} \int f(y_i|\{y_j\}_{j<i}, \Theta_s, \mathcal{M}_s)\pi(\Theta_s|\{y_j\}_{j<i}, \mathcal{M}_s)\, d\Theta_s. \end{aligned}$$

Therefore, the adequacy of the model as captured by the marginal likelihood corresponds to the cumulative out-of-sample predictive record where the fit of y_i is measured with respect to the posterior density using data $\{y_j\}_{j<i}$ up to the ith data point. This is in sharp contrast to in-sample measures of fit that condition on the entire data set y, or split-sample comparisons in which

the outcome may depend on the choice of estimation and comparison samples. By comparison, the marginal likelihood is invariant to permutations in the indices of the data, so that the same $m(y|\mathcal{M}_s)$ will be obtained if the data are arbitrarily rearranged.

Ivan: (1) Please elaborate on how marginal likelihood can be used for out of sample prediction (The above paragraph is short and hard to follow). (2) Maybe, also discuss how to evaluate out-of-sample prediction. One approach is scoring rules (Gneiting and Raftery, 2007), but I guess they are all arbitrary. This needs to be written carefully as the article by Gneiting and Raftery (2007) is well received in the literature (has more than 6000 citations). To tone it down, may be add that forecasting is not of paramount importance in our application. (3) Kindly also elaborate that comparisons across quantiles are not meaningful and that we should only compare a specific quantiles across models.

2.4.1 Marginal Likelihood of the GFREQ Model

To estimate the marginal likelihood of the FREQ model, we let $\Theta = (\beta, \Omega, \sigma)$ and employ the decomposition[5]

$$\pi(\beta^*, \Omega^*, \sigma^*|y) = \pi(\sigma^*|y)\, \pi(\beta^*|y, \sigma^*)\, \pi(\Omega^*|y, \beta^*, \sigma^*) \tag{11}$$

in the denominator of Equation (10), suppressing the model indicator $\mathcal{M}_s$ for notational simplicity. The posterior ordinate is evaluated at the posterior mean $(\beta^*, \Omega^*, \sigma^*)$ using the method of Chib and Jeliazkov (2001) since the conditional posterior for σ in Algorithm 3 is non-standard and sampling involves an MH step. The latent (α, γ, ν, h) are not included in the ordinate in Equation (11) to reduce dimensionality, computational costs, and estimation variability. Moreover, we place the ordinate $\pi(\sigma^*|y)$ first in Equation (11) to minimize MH sampling in the course of density estimation.

Following Chib and Jeliazkov (2001), we form an estimate of $\pi(\sigma^*|y)$ from the building blocks of Algorithm 3 as

$$\hat{\pi}(\sigma^*|y) = \frac{E_1\{\alpha_{MH}(\sigma, \sigma^*|y, \beta, \Omega, \alpha)\, q(\sigma, \sigma^*)|y, \beta, \Omega, \alpha)\}}{E_2\{\alpha_{MH}(\sigma^*, \sigma|y, \beta, \Omega, \alpha)\}}, \tag{12}$$

where the expectation E_1 in the numerator is with respect to the posterior $\pi(\sigma, \beta, \Omega, \alpha|y)$ and E_2 in the denominator is with respect to $\pi(\beta, \Omega, \alpha|y, \sigma^*)q(\sigma^*, \sigma|y)$. Therefore, draws from the *complete* MCMC run are used to form the average in the numerator of Equation (14), whereas draws from

[5]It is also possible to decompose equation 13 as $\pi(\beta^*, \Omega^*, \sigma^*|y) = \pi(\beta^*|y)\, \pi(\sigma^*, \Omega^*|y, \beta^*)$, however, we do not employ this decomposition to avoid the MH step for sampling σ in the reduced MCMC run.

a *reduced* MCMC run, in which Algorithm 1 is iterated conditionally on σ^* to produce draws from $\pi(\beta, \Omega, \alpha|y, \sigma^*)$, coupled with draws $\sigma \sim q(\sigma^*, \sigma|y, \beta, \Omega, \alpha)$, are employed in estimating the denominator.

We obtain $\pi(\beta^*|y, \sigma^*) = \int \pi(\beta^*|y, \sigma^*, \alpha, \Omega, \gamma, \nu, h)\pi(\alpha, \gamma, \Omega, \nu, h|y, \sigma^*)\, d\alpha\, d\gamma\, d\Omega\, d\nu\, dh$, in the same reduced run of Algorithm 1 in which σ is held at σ^*, so that $\hat{\pi}(\beta^*|y, \sigma^*) = E\{\pi(\beta^*|y, \sigma^*, \alpha, \gamma, \Omega, \nu, h)\}$, where the expectation is easily evaluated with those draws from $\pi(\alpha, \gamma, \Omega, \nu, h|y, \sigma^*)$.

To estimate $\pi(\Omega^*|y, \beta^*, \sigma^*) = \int \pi(\Omega^*|y, \beta^*, \sigma^*, \alpha, \gamma, \nu, h)\pi(\alpha, \gamma, \nu, h|y, \beta^*, \sigma^*)\, d\alpha\, d\gamma\, d\nu\, dh$, we conduct a second reduced MCMC run holding (β, σ) in Algorithm 1 fixed at (β^*, σ^*). The resulting draws from $\pi(\alpha, \gamma, \nu, h|y, \beta^*, \sigma^*)$ are used to produce $\hat{\pi}(\Omega^*|y, \beta^*, \sigma^*) = E\{\pi(\Omega^*|y, \beta^*, \sigma^*, \alpha, \gamma, \nu, h)\}$.

Combining the preceding three ordinate estimates through Equation (13) allows us to produce an estimate $\hat{\pi}(\beta^*, \Omega^*, \sigma^*|y)$ of the joint posterior ordinate for the denominator of Equation (10). The prior ordinates in the numerator of that expression are directly available. However, the likelihood ordinate requires integration of the density in Equation (3) over the random effects $\{\alpha_i\}$ and the shape parameters $\{\gamma_i\}$, i.e.,

$$f(y|\beta^*, \Omega^*, \sigma^*) = \prod_{i=1}^{n} \left[\int \left\{ \int \prod_{t=1}^{T_i} f_{GAL}\left(y_{it}|x_{it}'\beta^* + z_{it}'\alpha_i, \sigma^*, p_0, \gamma_i\right) d\gamma_i \right\} f(\alpha_i|\Omega^*) d\alpha_i \right],$$

where the integrals in the curly brackets are numerically computed using adaptive quadrature and the integrals in the square brackets are computed as averages of the terms in curly braces over draws $\alpha_i \sim f(\alpha_i|\Omega^*)$. Additionally, since $f_{GAL}(\cdot)$ is directly available from Equation (2), the mixture variables (ν, h) need not be involved in the likelihood computation.

2.4.2 Marginal Likelihood of the FREQ Model

To estimate the marginal likelihood of the FREQ model, we let $\Theta = (\beta, \Omega, \Theta_1)$ with $\Theta_1 = (\sigma, \gamma)$ and employ the decomposition

$$\pi(\beta^*, \Omega^*, \Theta_1^*|y) = \pi(\Theta_1^*|y)\, \pi(\beta^*|y, \Theta_1^*)\, \pi(\Omega^*|y, \beta^*, \Theta_1^*) \tag{13}$$

in the denominator of Equation (10), suppressing the model indicator $\mathcal{M}_s$ for notational simplicity. The posterior ordinate is evaluated at the posterior mean $(\beta^*, \Omega^*, \Theta_1^*)$ using the method of Chib and Jeliazkov (2001) since the conditional posterior for $\Theta_1 = (\sigma, \gamma)$ in Algorithm 1 is non-standard and sampling involves an MH step. The latent (α, ν, h) are not included in the ordinate in Equation (13) to reduce dimensionality, computational costs, and estimation variability. Moreover, we place the ordinate $\pi(\Theta_1^*|y)$ first in Equation (13) to minimize MH sampling in the course of density estimation.

Following Chib and Jeliazkov (2001), we form an estimate of $\pi(\Theta_1^*|y)$ from the building blocks of Algorithm 1 as

$$\hat{\pi}(\Theta_1^*|y) = \frac{E_1\{\alpha_{MH}(\Theta_1, \Theta_1^*|y, \beta, \Omega, \alpha)\, q(\Theta_1, \Theta_1^*)|y, \beta, \Omega, \alpha)\}}{E_2\{\alpha_{MH}(\Theta_1^*, \Theta_1|y, \beta, \Omega, \alpha)\}}, \tag{14}$$

where the expectation E_1 in the numerator is with respect to the posterior $\pi(\Theta_1, \beta, \Omega, \alpha|y)$ and E_2 in the denominator is with respect to $\pi(\beta, \Omega, \alpha|y, \Theta_1^*)q(\Theta_1^*, \Theta_1|y)$. Therefore, draws from the *complete* MCMC run are used to form the average in the numerator of Equation (14), whereas draws from a *reduced* MCMC run, in which Algorithm 1 is iterated conditionally on Θ_1^* to produce draws from $\pi(\beta, \Omega, \alpha|y, \Theta_1^*)$, coupled with draws $\Theta_1 \sim q(\Theta_1^*, \Theta_1|y, \beta, \Omega, \alpha)$, are employed in estimating the denominator.

We obtain $\pi(\beta^*|y, \Theta_1^*) = \int \pi(\beta^*|y, \Theta_1^*, \alpha, \Omega, \nu, h)\pi(\alpha, \Omega, \nu, h|y, \Theta_1^*)\, d\alpha\, d\Omega\, d\nu\, dh$, in the same reduced run of Algorithm 1 in which Θ_1 is held at Θ_1^*, so that $\hat{\pi}(\beta^*|y, \Theta_1^*) = E\{\pi(\beta^*|y, \Theta_1^*, \alpha, \Omega, \nu, h)\}$, where the expectation is easily evaluated with those draws from $\pi(\alpha, \Omega, \nu, h|y, \Theta_1^*)$.

To estimate $\pi(\Omega^*|y, \beta^*, \Theta_1^*) = \int \pi(\Omega^*|y, \beta^*, \Theta_1^*, \alpha, \nu, h)\pi(\alpha, \nu, h|y, \beta^*, \Theta_1^*)\, d\alpha\, d\nu\, dh$, we conduct a second reduced MCMC run holding (β, Θ_1) in Algorithm 1 fixed at (β^*, Θ_1^*). The resulting draws from $\pi(\alpha, \nu, h|y, \beta^*, \Theta_1^*)$ are used to produce $\hat{\pi}(\Omega^*|y, \beta^*, \Theta_1^*) = E\{\pi(\Omega^*|y, \beta^*, \Theta_1^*, \alpha, \nu, h)\}$.

Combining the preceding three ordinate estimates through Equation (13) allows us to produce an estimate $\hat{\pi}(\beta^*, \Omega^*, \Theta_1^*|y)$ of the joint posterior ordinate for the denominator of Equation (10). The prior ordinates in the numerator of that expression are directly available. However, the likelihood ordinate requires integration of the density in Equation (3) over the random effects $\{\alpha_i\}$, i.e.,

$$f(y|\beta^*, \Omega^*, \sigma^*, \gamma^*) = \prod_{i=1}^{n} \left[\int \left\{ \prod_{t=1}^{T_i} f_{GAL}\left(y_{it}|x_{it}'\beta^* + z_{it}'\alpha_i, \sigma^*, p_0, \gamma^*\right) \right\} f(\alpha_i|\Omega^*) d\alpha_i \right],$$

where the integrals in the square brackets are computed as averages of the terms in curly braces over draws $\alpha_i \sim f(\alpha_i|\Omega^*)$. Additionally, since $f_{GAL}(\cdot)$ is directly available from Equation (2), the mixture variables (ν, h) need not be involved in the likelihood computation.

The marginal likelihood of the GFREQ model follows the same steps as in section 2.4.2 by setting $\Theta_1 = \sigma$ and treating $\{\gamma_i\}$ as latent parameters. [Ivan, Shubham: please describe the computation of ML for the GFREQ model in a separate section.]

2.4.3 Marginal Likelihood of the REQ Model

Because all conditional densities in Algorithm 2 are of known form, we compute the marginal likelihood of the REQ model using the method of Chib (1995). We proceed analogously to the case in Section 2.4.2 and decompose the joint posterior ordinate (marginally of α and ν) as

$$\pi(\beta^*, \sigma^*, \Omega^*|y) = \pi(\beta^*|y)\, \pi(\Omega^*|y, \beta^*)\, \pi(\sigma^*|y, \beta^*, \Omega^*). \tag{15}$$

The posterior density $\pi(\beta^*|y) = \int \pi(\beta^*|y, \nu, \sigma, \Omega)\, \pi(\nu, \sigma, \Omega|y)\, d\nu\, d\sigma\, d\Omega$ is estimated as the expectation $\hat{\pi}(\beta^*|y) = E\{\pi(\beta^*|y, \nu, \sigma, \Omega)\}$ with draws from $\pi(\nu, \sigma, \Omega|y)$ obtained in the main run. To estimate $\pi(\Omega^*|y, \beta^*) = \int \pi(\Omega^*|y, \beta^*, \sigma, \alpha, \nu)\, \pi(\sigma, \alpha, \nu|y, \beta^*)\, d\sigma\, d\alpha\, d\nu$, we form a reduced run of Algorithm 2, where β is set to β^*, to simulate draws from $\pi(\sigma, \alpha, \nu|y, \beta^*)$ that are used to produce $\hat{\pi}(\Omega^*|y, \beta^*) = E\{\pi(\Omega^*|y, \beta^*, \sigma, \alpha, \nu)\}$. Finally, a second reduced run of Algorithm 2 with (β, Ω) fixed at (β^*, Ω^*) is used to estimate $\pi(\sigma^*|y, \beta^*, \Omega^*) = \int \pi(\sigma^*|y, \beta^*, \Omega^*, \alpha, \nu)\, \pi(\alpha, \nu|y, \beta^*, \Omega^*)\, d\alpha\, d\nu$ as $\hat{\pi}(\sigma^*|y, \beta^*, \Omega^*) = E\{\pi(\sigma^*|y, \beta^*, \Omega^*, \alpha, \nu)\}$, where expectation is with respect to $\pi(\alpha, \nu|y, \beta^*, \Omega^*)$.

As before, the prior ordinates are readily available since the prior distributions for (β, σ, Ω) have tractable forms, but the likelihood requires integration over of the latent $\{\alpha_i\}$,

$$f(y|\beta^*, \sigma^*, \Omega^*) = \prod_{i=1}^{n} \left[\int \left\{ \prod_{t=1}^{T_i} f_{AL}\left(y_{it}|x'_{it}\beta^* + z'_{it}\alpha_i, \sigma^*, p\right) \right\} f(\alpha_i|\Omega^*)\, d\alpha_i \right],$$

which is easily obtained by simulation. By virtue of using the AL density directly, rather than its mixture representation, the latent ν need not be involved in the computations.

3 Simulation Studies

[POINTS TO CONSIDER: To lighten the simulation study section and address #3 from Reviewer 1, we can do the following:

(1) In the simulation study, present the results from the GFREQ model only for SS9 along with a box-plot of the posterior mean of γ_i's.

(2) Shift the results from FREQ models for all nine simulation studies and the trace plots to the appendix.]

To illustrate the performance of the proposed modeling and estimation techniques, we conduct a series of simulations. Moreover, we compute marginal likelihoods to examine the relative performance of the FREQ model *vis-a-vis* the REQ model. The data for the simulation studies are generated from the panel data model,

$$y_{it} = \alpha_{1i} + \alpha_{2i}\, z_{2it} + \beta_1 + \beta_2\, x_{2it} + \beta_3\, x_{3it} + \varepsilon_{it},$$

where $\alpha = (\alpha_1, \alpha_2)' \sim N_2([0,0]', [1,0;0,1])$, $\beta = (\beta_1, \beta_2, \beta_3)' = (10, 5, 2)'$, $z_{2it} \sim \text{Unif}(1,3)$, $x_2 \sim N(0, 0.25)$, $x_3 \sim N(2, 0.25)$, and the errors ε were generated from a standard logistic distribution $\mathcal{L}(0,1)$. We generate 9 different data samples with $T_i = (5, 10, 15)$ and $n = (100, 250, 500)$, where T_i denotes the number of repeated observations for each individual i and n represents the number of individuals.

In each simulation study, the posterior estimates of the parameters in the FREQ model are obtained based on the simulated data and the following prior distributions: $\beta \sim N(0_k, 100 I_k)$, $\Omega \sim IW(r_0, R_0)$, $\sigma \sim IG(5/2, 8/2)$, and $\gamma \sim \text{Unif}(L, U)$. Here, $r_0 = 5 + l$, $R_0 = (r_0 - l - 1) * I_l$, and (L, U) are obtained as mentioned in Section 2. The same prior distributions are employed for the REQ model. Table 1 reports, for each simulated data set, the MCMC results at five different quantiles obtained from 10,000 iterations after a burn-in of 2,500 iterations[6] Inefficiency factors are calculated as $1 + 2\sum_{t=1}^{T} \rho_k(t)\,((T-t)/T)$, where $\rho_k(t)$ denotes the autocorrelation for the kth parameter at lag t, and T is the value at which the autocorrelations taper off. In the MH sampling of (σ, γ), the tuning factor ι is adjusted to obtain an acceptance rate of approximately 0.3. Convergence of the MCMC draws is quick – a fairly representative case is displayed in Figure 2, with the trace plots across quantiles in our 9 simulated samples exhibiting similar behavior. For all cases that were considered, we report inefficiency factors in Table 1.

Table 1: Posterior mean (MEAN), standard deviation (SD) and inefficiency factor (IF) of the parameters in the family of FREQ models from nine simulation studies: SS1 ($T_i = 5, n = 100$), SS2 ($T_i = 5, n = 250$), SS3 ($T_i = 5, n = 500$), SS4 ($T_i = 10, n = 250$), SS5 ($T_i = 10, n = 250$), SS6 ($T_i = 10, n = 500$), SS7 ($T_i = 15, n = 100$), SS8 ($T_i = 15, n = 250$, and SS9 ($T_i = 15, n = 500$))

	10TH QTL			25TH QTL			50TH QTL			75TH QTL			90TH QTL		
SS1	MEAN	SD	IF	MEAN	SD	IF	MEAN	SD	IF	MEAN	SD	IF	MEAN	SD	IF
β_1	7.42	0.73	4.81	8.53	0.73	3.15	9.40	0.74	2.48	10.57	0.74	3.67	11.63	0.75	5.36
β_2	5.11	0.35	4.97	5.17	0.34	3.18	5.24	0.34	2.27	5.13	0.35	3.81	4.96	0.37	5.77

[6] In general, the computation time required for the FREQ model is approximately two to three times longer than that of the REQ model. For instance, on a MacBook Pro equipped with an Apple M1 chip and 16GB of RAM, the SS1 (SS9) simulation exercise at the 10th quantile takes 614 seconds (1 hour and 18 minutes) to estimate and compute the marginal likelihood for the FREQ model, whereas the REQ model requires only 288 seconds (27 minutes). The computation time for GFREQ on a AMD Ryzen 16-core processor with 128GB of RAM for SS1 (SS9) were 1 hour 16 minutes (2 hours 32 minutes).

Table 1 – *Continued from previous page*

β_3	2.42	0.34	4.93	2.34	0.34	3.24	2.37	0.34	2.36	2.31	0.35	3.76	2.28	0.36	5.75
σ	0.44	0.03	9.39	0.57	0.05	11.22	0.65	0.03	8.50	0.53	0.06	10.79	0.44	0.03	7.02
γ	2.82	0.32	9.40	1.11	0.21	10.69	−0.10	0.09	13.46	−1.28	0.21	11.04	−2.92	0.28	7.38
Ω_{11}	1.08	0.49	15.25	0.93	0.42	14.82	1.01	0.47	16.13	0.85	0.39	15.33	0.73	0.37	16.37
Ω_{22}	0.98	0.21	11.09	1.03	0.22	10.94	0.99	0.22	9.18	1.05	0.23	12.27	1.15	0.25	14.81
Ω_{12}	0.36	0.22	6.82	0.30	0.21	6.54	0.34	0.22	6.70	0.28	0.22	8.68	0.15	0.24	8.53
SS2	10th qtl			25th qtl			50th qtl			75th qtl			90th qtl		
β_1	7.79	0.47	6.43	8.92	0.47	4.07	9.90	0.46	2.49	10.99	0.46	4.51	12.16	0.46	5.62
β_2	5.12	0.22	5.51	5.14	0.21	3.67	5.16	0.21	2.26	5.18	0.21	4.25	5.22	0.22	5.71
β_3	2.14	0.22	6.47	2.04	0.22	4.10	2.04	0.21	2.31	2.00	0.22	4.62	1.88	0.22	5.78
σ	0.42	0.02	9.14	0.50	0.06	14.79	0.64	0.02	11.08	0.47	0.06	13.05	0.41	0.02	11.02
γ	2.99	0.21	8.87	1.34	0.19	14.60	0.00	0.07	13.13	−1.45	0.19	12.80	−3.02	0.21	9.74
Ω_{11}	0.73	0.33	37.12	0.81	0.37	28.74	1.01	0.45	35.36	0.70	0.31	39.75	0.91	0.43	43.20
Ω_{22}	1.17	0.18	27.03	1.12	0.17	24.05	1.15	0.18	29.03	1.11	0.17	24.00	1.15	0.19	46.11
Ω_{12}	−0.01	0.21	12.15	0.01	0.21	12.20	−0.05	0.23	11.31	0.05	0.19	10.84	−0.03	0.24	15.94
SS3	10th qtl			25th qtl			50th qtl			75th qtl			90th qtl		
β_1	8.08	0.33	5.20	9.01	0.34	4.63	10.16	0.33	2.13	11.18	0.32	3.80	12.05	0.34	5.56
β_2	5.07	0.16	5.78	5.08	0.16	4.85	5.11	0.15	2.23	5.08	0.16	4.28	5.03	0.16	5.88
β_3	1.98	0.16	5.64	1.98	0.16	4.44	1.91	0.15	2.17	1.91	0.15	4.15	1.95	0.16	5.79
σ	0.42	0.02	12.82	0.46	0.07	32.09	0.66	0.02	10.72	0.50	0.05	20.87	0.42	0.02	11.01
γ	3.03	0.16	10.36	1.52	0.19	30.61	−0.02	0.05	15.49	−1.42	0.16	18.99	−3.01	0.16	9.53
Ω_{11}	1.59	0.55	84.89	0.95	0.43	76.74	1.53	0.52	42.68	1.12	0.45	77.29	1.27	0.50	75.02
Ω_{22}	1.10	0.14	93.97	1.02	0.12	66.65	1.11	0.13	30.89	1.04	0.13	52.21	1.09	0.13	45.04
Ω_{12}	−0.30	0.24	28.82	−0.09	0.19	24.79	−0.30	0.22	12.06	−0.15	0.20	20.99	−0.23	0.22	19.94
SS4	10th qtl			25th qtl			50th qtl			75th qtl			90th qtl		
β_1	8.81	0.47	5.40	9.68	0.48	4.04	10.54	0.47	2.62	11.59	0.47	4.62	12.80	0.52	6.77
β_2	4.99	0.22	6.22	5.00	0.23	3.83	5.00	0.22	2.66	4.96	0.22	4.58	5.01	0.24	6.75
β_3	1.77	0.22	5.53	1.78	0.22	4.16	1.83	0.22	2.70	1.81	0.22	4.78	1.71	0.24	7.04
σ	0.43	0.02	8.41	0.55	0.04	9.57	0.65	0.02	6.29	0.49	0.05	9.12	0.42	0.02	8.57
γ	2.88	0.21	7.78	1.19	0.14	10.23	−0.02	0.07	12.45	−1.40	0.15	8.55	−2.93	0.18	7.15
Ω_{11}	0.86	0.36	17.14	0.78	0.30	11.05	0.88	0.35	13.78	0.68	0.27	11.07	0.97	0.42	17.73
Ω_{22}	1.07	0.19	12.67	1.02	0.18	7.77	1.04	0.19	7.78	1.03	0.18	8.29	1.06	0.20	13.51
Ω_{12}	0.15	0.20	4.09	0.20	0.18	3.17	0.16	0.18	2.99	0.20	0.17	4.02	0.10	0.22	4.82
SS5	10th qtl			25th qtl			50th qtl			75th qtl			90th qtl		
β_1	7.99	0.33	5.94	9.12	0.32	4.68	10.29	0.32	2.59	11.24	0.31	4.15	12.22	0.31	6.02
β_2	4.79	0.15	7.29	4.77	0.15	4.40	4.75	0.14	2.67	4.80	0.15	4.16	4.83	0.15	7.82
β_3	1.98	0.15	6.27	1.88	0.15	4.81	1.81	0.15	2.66	1.85	0.15	4.29	1.84	0.15	6.35
σ	0.43	0.01	7.55	0.52	0.04	18.91	0.67	0.01	5.80	0.54	0.04	13.58	0.44	0.01	7.40
γ	2.94	0.13	6.35	1.36	0.14	17.66	0.01	0.04	12.62	−1.30	0.12	13.14	−2.92	0.12	7.32
Ω_{11}	1.14	0.35	31.26	0.96	0.29	21.63	1.37	0.34	11.67	0.91	0.29	21.99	0.97	0.31	25.82
Ω_{22}	0.87	0.11	21.20	0.83	0.11	16.79	0.89	0.11	8.46	0.85	0.11	13.62	0.87	0.11	17.06
Ω_{12}	−0.00	0.15	6.18	0.08	0.13	6.22	−0.08	0.15	3.42	0.07	0.13	5.53	0.06	0.14	7.35
SS6	10th qtl			25th qtl			50th qtl			75th qtl			90th qtl		
β_1	8.17	0.23	6.25	9.18	0.22	4.71	10.14	0.21	2.45	11.15	0.22	4.21	12.20	0.23	6.89
β_2	5.08	0.11	6.34	5.06	0.10	4.85	5.06	0.10	2.69	5.08	0.10	4.42	5.11	0.11	6.99

Table 1 – *Continued from previous page*

β_3	1.97	0.10	6.88	1.92	0.10	4.91	1.95	0.10	2.56	1.95	0.10	4.36	1.89	0.10	7.35
σ	0.42	0.01	8.02	0.50	0.04	38.95	0.66	0.01	6.18	0.53	0.03	17.79	0.42	0.01	8.39
γ	2.96	0.10	7.99	1.39	0.13	35.65	0.01	0.03	10.70	−1.29	0.09	16.57	−2.96	0.10	7.94
Ω_{11}	1.39	0.30	33.79	1.02	0.24	30.53	1.26	0.25	17.14	0.92	0.22	24.94	1.41	0.30	32.57
Ω_{22}	0.95	0.09	21.46	0.92	0.08	20.39	0.96	0.09	12.01	0.92	0.08	18.93	0.99	0.09	27.75
Ω_{12}	−0.09	0.13	9.49	0.02	0.11	6.76	−0.08	0.11	4.07	0.03	0.11	6.17	−0.13	0.13	8.39
SS7	10th qtl			25th qtl			50th qtl			75th qtl			90th qtl		
β_1	7.49	0.39	6.12	8.58	0.40	5.73	9.72	0.40	2.66	10.64	0.40	5.16	11.66	0.40	7.22
β_2	5.18	0.18	7.18	5.15	0.18	6.36	5.05	0.17	2.80	5.15	0.18	5.43	5.20	0.18	7.59
β_3	2.25	0.18	7.31	2.22	0.18	6.33	2.14	0.18	2.79	2.19	0.19	5.78	2.19	0.18	7.94
σ	0.42	0.02	7.50	0.46	0.05	12.29	0.66	0.02	5.35	0.49	0.06	12.40	0.43	0.02	7.96
γ	3.00	0.16	7.06	1.52	0.16	11.36	−0.01	0.05	11.77	−1.41	0.19	12.32	−2.97	0.17	8.17
Ω_{11}	1.39	0.45	15.70	1.01	0.34	12.25	1.08	0.34	8.91	1.07	0.36	13.10	1.03	0.36	15.25
Ω_{22}	0.85	0.15	7.46	0.85	0.15	6.23	0.88	0.15	3.99	0.86	0.15	6.67	0.87	0.15	8.27
Ω_{12}	0.25	0.18	4.37	0.31	0.15	3.18	0.27	0.15	2.19	0.28	0.16	2.91	0.25	0.16	3.56
SS8	10th qtl			25th qtl			50th qtl			75th qtl			90th qtl		
β_1	7.80	0.27	8.34	9.00	0.26	4.24	10.14	0.25	2.68	11.11	0.26	4.98	11.98	0.27	6.98
β_2	4.95	0.12	6.29	4.96	0.12	4.37	4.95	0.12	3.01	4.95	0.12	5.43	4.95	0.13	8.17
β_3	2.02	0.12	8.23	1.92	0.12	4.42	1.87	0.11	2.77	1.91	0.12	5.27	1.99	0.12	7.52
σ	0.45	0.01	7.26	0.57	0.02	9.27	0.68	0.01	5.04	0.51	0.04	12.24	0.44	0.01	7.83
γ	2.89	0.11	7.10	1.22	0.08	9.41	−0.05	0.03	11.58	−1.43	0.11	12.24	−2.95	0.10	7.53
Ω_{11}	0.99	0.26	18.59	0.91	0.23	13.74	1.10	0.24	9.19	0.89	0.24	17.06	1.00	0.25	15.75
Ω_{22}	1.09	0.12	12.74	1.02	0.11	9.58	1.04	0.12	8.09	1.02	0.11	13.22	0.98	0.11	11.77
Ω_{12}	0.01	0.13	3.80	0.11	0.12	3.60	0.07	0.13	3.10	0.11	0.12	3.68	0.11	0.13	4.41
SS9	10th qtl			25th qtl			50th qtl			75th qtl			90th qtl		
β_1	8.10	0.19	6.51	9.17	0.18	3.92	10.21	0.18	2.66	11.25	0.18	4.43	12.26	0.19	6.89
β_2	5.06	0.09	6.69	5.06	0.08	4.26	5.09	0.08	3.01	5.05	0.08	4.53	5.04	0.09	7.14
β_3	1.97	0.09	7.11	1.94	0.08	4.41	1.94	0.08	2.77	1.94	0.08	4.54	1.92	0.09	7.29
σ	0.45	0.01	6.84	0.58	0.01	6.41	0.69	0.01	6.17	0.56	0.02	7.00	0.45	0.01	6.87
γ	2.89	0.07	6.60	1.20	0.05	8.20	−0.00	0.02	8.79	−1.28	0.06	7.30	−2.89	0.07	6.46
Ω_{11}	0.81	0.18	26.07	0.79	0.17	20.56	1.01	0.16	11.33	0.73	0.16	19.95	0.92	0.20	24.38
Ω_{22}	1.12	0.09	16.91	1.11	0.09	11.11	1.16	0.09	6.08	1.12	0.09	11.78	1.12	0.09	17.18
Ω_{12}	−0.02	0.10	4.45	−0.02	0.09	3.18	−0.13	0.09	2.35	−0.02	0.09	3.78	−0.04	0.10	3.91

The results, presented in Table 1, show that the regression coefficients β are estimated well with intercepts being quantile dependent, and that inefficiency factors are low, indicating good mixing of the proposed MCMC sampler across quantiles and simulation studies. For the scale parameter σ, the posterior estimates vary with quantiles, have small standard deviations, and low inefficiency factors. The posterior estimates of the components of covariance matrix (i.e., Ω_{11}, Ω_{22}, and Ω_{12}) are also close to the true values used to generate the data. However, inefficiency factors for these parameters tend to be higher as compared to other model parameters indicating slower

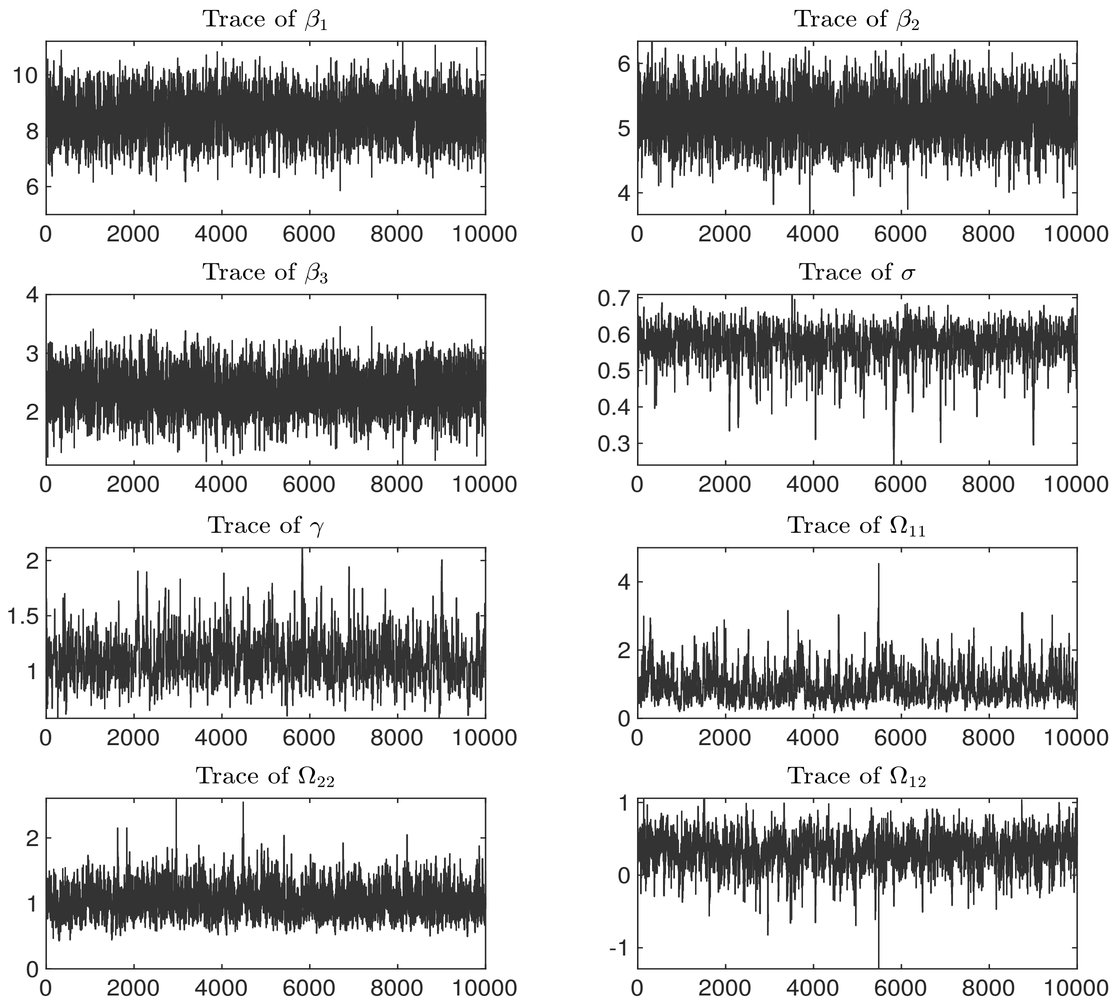

Figure 2: Trace plots from FREQ model, Simulation Study 1 ($m = 5$, $n = 100$), 25th quantile.

mixing. This is not unusual, however, for parameters that are further removed from the data y in hierarchical models (here Ω depends on the data only via $\{\alpha_i\}$) and has been reported elsewhere in the literature (Chib and Jeliazkov, 2006).

The posterior estimates of the shape parameter γ suggest that the GAL distribution allows considerable flexibility in skewness while modeling the data relative to the AL distribution. For example, in Simulation Study 9, the posterior mean of γ at the 25th (75th) quantile is 1.20 (-1.28) which corresponds to a skewness of -0.06 (0.13). This is in contrast to the fixed skewness of 1.64

Table 2: Logarithm of marginal likelihoods for the GFREQ, FREQ, and REQ models in 9 simulations studies.

	10TH QTL	25TH QTL	50TH QTL	75TH QTL	90TH QTL
SS1-GFREQ	−994	−1061	−1100	−1065	−1002
SS1-FREQ	−1157	−1151	−1153	−1152	−1164
SS1-REQ	−1209	−1172	−1153	−1185	−1232
SS2-GFREQ	−2436	−2608	−2704	−2606	−2430
SS2-FREQ	−2832	−2818	−2828	−2817	−2834
SS2-REQ	−2955	−2883	−2828	−2885	−2952
SS3-GFREQ	−4870	−5225	−5246	−5235	−4887
SS3-FREQ	−5681	−5656	−5679	−5660	−5688
SS3-REQ	−5885	−5771	−5677	−5785	−5919
SS4-GFREQ	−2034	−2088	−2126	−2090	−2048
SS4-FREQ	−2171	−2159	−2168	−2157	−2182
SS4-REQ	−2314	−2216	−2168	−2234	−2342
SS5-GFREQ	−5139	−5225	−5354	−5256	−5137
SS5-FREQ	−5469	−5421	−5448	−5424	−5468
SS5-REQ	−5863	−5591	−5445	−5592	−5846
SS6-GFREQ	−10197	−10427	−10621	−10429	−10188
SS6-FREQ	−10857	−10759	−10803	−10764	−10847
SS6-REQ	−11635	−11105	−10802	−11089	−11583
SS7-GFREQ	−3105	−3132	−3173	−3134	−3096
SS7-FREQ	−3212	−3180	−3201	−3185	−3206
SS7-REQ	−3490	−3296	−3200	−3291	−3471
SS8-GFREQ	−7822	−7907	−8008	−7917	−7858
SS8-FREQ	−8114	−8054	−8088	−8052	−8141
SS8-REQ	−8748	−8300	−8089	−8357	−8864
SS9-GFREQ	−15764	−15895	−16079	−15889	−15789
SS9-FREQ	−16322	−16172	−16222	−16161	−16354
SS9-REQ	−17728	−16726	−16220	−16758	−17807

and −1.64 in the REQ model. Additionally, the posterior mean at the 50th quantile is −0.00 with a high standard deviation (0.02), which implies zero skewness. In summary, all estimates of γ suggest that skewness corresponding to the FREQ model (vis-a-vis the REQ model) is much closer to that of the underlying data generating process.

While the additional flexibility of the FREQ model *vis-a-vis* the REQ model can be established both theoretically and in the simulations considered here, a pertinent research question is what can be gained from this flexibility and whether it translates to a better model fit that may justify the modeling and simulation costs. To address this issue, we report the log-marginal likelihoods for the two models in Table 2. The results in the table lend evidence to two facts that are broadly supported across samples. The first is that the gains from flexibility increase as we move towards the tail of the distribution – the data favor the FREQ model over the REQ model more strongly at more extreme quantiles. Suggested text: The second is that although Bayesian model comparisons penalize model complexity, the FREQ model remains competitive with the REQ model under most settings even at the median (i.e., 50th quantile). Because the assumed error distribution in our simulations is symmetric, at the 50th quantile the more profligate FREQ modeling of skewness is not warranted – Table 1 reveals that the estimates of γ are practically zero, indicating a lack of skewness and an equivalence between FREQ and REQ models. However, Table 2 shows that logarithm of marginal likelihoods are identical in three settings, the FREQ and REQ are comparable in 3 settings ($1 \leq BF_{\frac{REQ}{FREQ}} \leq 3.2$), REQ has substantial evidence in 2 settings ($3.2 \leq BF_{\frac{REQ}{FREQ}} \leq 10$), and REQ has strong evidence under 1 setting ($10 \leq BF_{\frac{REQ}{FREQ}} \leq 100$). We therefore expect that the flexibility provided by FREQ modeling will outperform the REQ modeling, almost always, but a formal model comparison will be extremely useful for any given application, especially for asymmetric densities or tail analysis. We employ these considerations in our application in the next section

The second is that Bayesian model comparisons penalize model complexity in settings where it is unnecessary. Because the assumed error distribution in our simulations is symmetric, at the 50th quantile the more profligate FREQ modeling of skewness is not warranted – Table 1 reveals that the estimates of γ are practically zero, indicating a lack of skewness and an equivalence between FREQ and REQ, whereas Table 2 shows that marginal likelihoods unequivocally prefer the more parsimonious REQ model given that the extra skewness parameter is of little value in the absence of skewness. We therefore expect that the appropriateness of AL versus GAL modeling will be application specific and should be accompanied by formal model comparisons, especially for asymmetric densities or tail analysis. We employ these considerations in our application in the next section.

4 Application to U.S. Residential Rents

The US housing sector has received considerable attention as a result of the Global Financial Crisis (GFC) when mortgage delinquencies and foreclosures increased. Homeownership rates fell from 69% in 2004 to 62% in 2016. Because of the nexus between house prices and residential rents (see Loewenstein and Willen (2023) for a recent study), a drop in homeownership is likely to increase demand for rental units. However, the GFC also featured a drop in household income and an increase in unemployment.[7] The concurrence of these events, theoretically, leads to ambiguous effects on rental rates. In this application, we take an empirical approach and explore changes to residential rental rates in the post-GFC United States. In particular, we examine how median rental prices in 14,533 zip codes are influenced by unemployment rates and mortgage policies.

In studying rental rates, three concerns must be addressed: (1) heterogeneity because regions of the U.S. vary greatly by housing supply, income, and other factors that influence prices, (2) skewness in the distribution of prices because that distribution has a large right tail and significant outliers, and (3) heterogeneity in the skewness parameter to allow for increased flexibility in shape of the error distribution. Exhibiting some of these concerns, Figure 3 presents box plots of median rental rates in 5 states and the entire U.S. for 2010 and 2016. Apparent from the figure are the dramatic differences across the states, where California's lower quartile is above most other states' upper quartile. Additionally, all states display a strong upper (right) skew. Thus, we employ our GFREQ and FREQ models to accommodate heterogeneity and allow for skewness flexibility in the error term. The performance of the GFREQ and FREQ models are tested relative to the REQ model using our novel marginal likelihood approach. This exercise provides important insights for understanding how the data support the different specifications across various quantiles.

There is a vast literature on house prices and rental rates both before and after the GFC. Studies have examined various price determinants including zoning, regulation, and housing supply (Glaeser et al., 2020; Jackson, 2018), income differences (Quigley and Raphael, 2004), and tax policy (Chatterjee and Eyigungor, 2015). Many of these studies highlight how heterogeneity in city-specific features, such as average income and land availability, can lead to discrepancies in the effect of the boom and bust on prices. Additionally, a few international papers have focused on

[7] Aggregate figures for mortgage delinquencies, unemployment, household income, and homeownership are available on FRED.

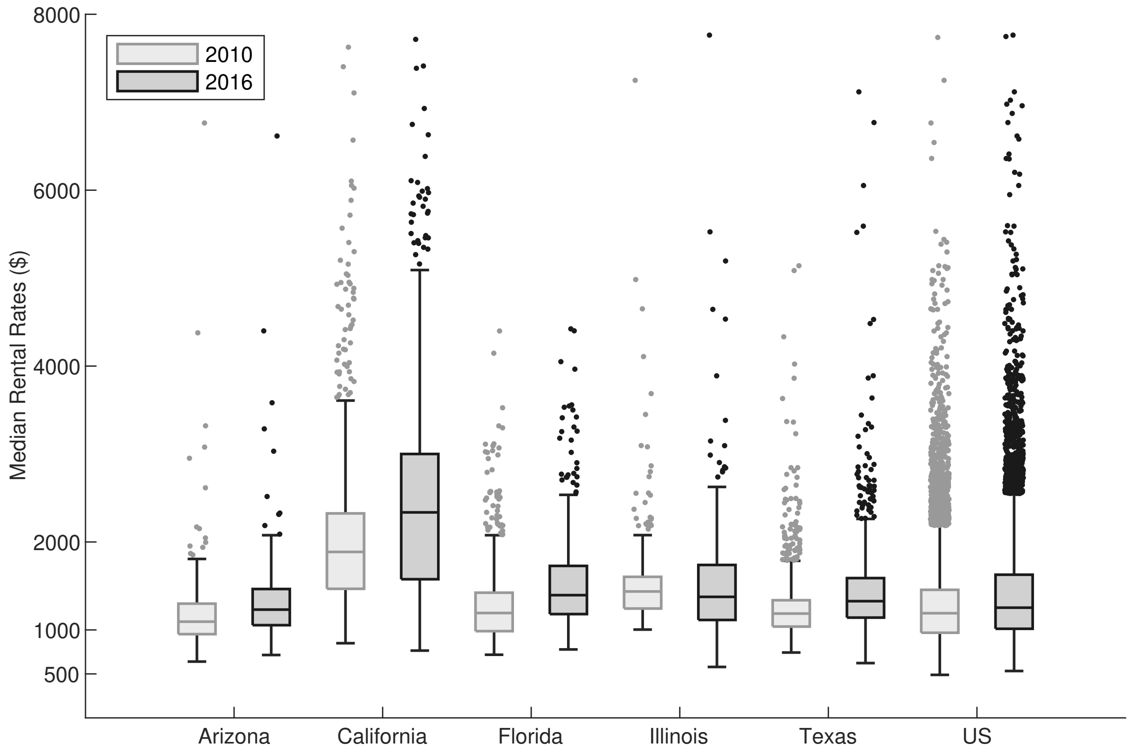

Figure 3: Box plots of median rental rates in 5 states and the entire US for 2010 and 2016. Note that the y-axis has been capped at $8000 for better visual representation.

the distribution or quantiles of rental prices (Thomschke, 2015; März et al., 2016; Waltl, 2018). We contribute to this literature by implementing a novel quantile regression approach to study rental rates in the United States. Moving beyond mean regression and exploiting large differences in population, income, and economic activity provides a deeper understanding of the determinants of rental markets. Further, our new methodology and model comparison approaches allow us to uncover potential biases that may result from ignored heterogeneity and erroneous distributional assumptions.

4.1 Data

We construct a novel zip-code-level data set where our outcome variable of interest, y_{it}, is the median monthly rental price of zip code i at year t. The sample includes $n = 14,533$ zip codes in the United States from 2010-2016 ($T = 7$). The residential rental price data come from the Zillow Rental Index (ZRENT). Our covariates include annual controls for each zip code's population, demographics, socioeconomic status, agriculture, property ownership, mortgage characteristics, and unemployment. The covariates are constructed from the Statistics of Income (SOI) Tax Stats, Individual Income Tax Statistics, provided by the Internal Revenue Service (IRS). Table 3 presents descriptions and summary statistics of our variables. We proxy for "population" using the total number of tax returns filed in the zip code and the remaining variables are generally a function of that measure.

Table 3: Data summary.

VARIABLE	Description	Mean	SD	Max	Min
LnRent	Logarithm of median monthly rental price	7.177	0.378	9.745	6.082
SSBfrac	Fraction of the population receiving social security benefits	0.139	0.059	0.814	0
Farmfrac	Fraction of the population receiving farming credits	0.019	0.033	0.315	0
REfrac	Fraction of the population with real estate taxes	0.273	0.134	0.821	0
HMRate	Fraction of the population with home mortgage deductions	0.236	0.116	0.786	0
AltMinRate	Fraction of the population paying alternative minimum taxes	0.027	0.050	0.444	0
EnergyRate	Fraction of the population receiving energy tax credits	0.024	0.019	0.147	0
EITCRate	Fraction of the population receiving earned income tax credits	0.184	0.099	0.711	0
UnempRate	Fraction of the population receiving unemployment compensation	0.072	0.043	0.518	0.001
lAvgAGI	ln-average adjusted gross income	4.040	0.458	7.899	1.610
lreturn	ln-number of returns filed (proxy for population)	8.491	1.103	10.902	4.605

We model the data using the GFREQ, FREQ, and REQ models, where $y_{it} = LnRent_{it}$, x_{it} includes the remaining variables in Table 3, and z_{it} is a constant to control for zip-code-level heterogeneity. Heterogeneity is an important concern. While a researcher can control for a host of demographic, socioeconomic, and location characteristics, much is left unobserved. City-level policies, nearby neighborhood spillovers, and commuting effects may enter the error term, which heavily influence rental prices. Thus, our specifications for all models include zip code random effects and the GFREQ model allows for heterogeneity in the shape of the error distribution. Lastly, we include time indicators to capture aggregate changes to prices.

4.2 Training Sample Priors

Prior distributions play an important role in Bayesian inference, particularly in model comparison where marginal likelihoods and hence their ratios, the Bayes factors, become arbitrary with improper priors, or sensitive to the prior with formally proper, but increasingly diffuse priors. Therefore, we employ a training sample approach where we take 10% of our data as a training sample and retain the remainder as a comparison sample. The data in the training sample are used to construct a first-stage posterior distribution which is used as a proper informative training sample prior when analyzing the comparison sample. Information from the training sample is not lost, as it is now part of the prior density used in the analysis of the remaining 90% of the data.[8]

4.3 Results

Before getting to the parameter estimates, we bring attention to the marginal likelihood results. We compare the performance of the GFREQ and FREQ models, relative to the REQ model, for the full sample of U.S. zip codes, as well as several state-specific models (Arizona, California, and Illinois). We consider these smaller samples of states to empirically explore the model fit when n is smaller and when there are varying degrees of heterogeneity and skewness in the sample. Table 4 presents the log-marginal likelihood estimates for the four samples across five quantiles. We find that the GFREQ and FREQ models have a higher marginal likelihood than the REQ model in all samples at the 10th, 25th, 50th, 75th, and 90th quantiles. The GFREQ model clearly dominates across all considerations. The differences, particularly further in the tails, are quite dramatic, giving the GFREQ model a posterior model probability of ≈ 1 over the other models. The dominance of the GFREQ is also most apparent in the more heterogeneous samples of California and US, where zip codes have vast differences in populations and incomes. Figures 4 and 5 provide boxplots of the estimates of the γ_is for two of our samples – the entire US and Illinois – which clearly demonstrate the need for flexibility in the parameter. Allowing for heterogeneity in the shape of the error distribution over zip codes i has clear advantages. Overall, these results demonstrate strong support from the data for the additional flexibility of the FREQ and GFREQ models. Researchers especially interested in the tail of their distribution of interest should employ these

[8]When estimating the model on the training sample, Arizona, California and Illinois, we used the following weak priors: $\beta \sim N(0_k, 25I_k)$, $\varphi^2 \sim IG(12/2, 10/2)$, $\sigma \sim IG(10/2, 8/2)$ and $\gamma \sim \text{Unif}(L, U)$, where (L, U) are obtained as in Section 2.

Table 4: Logarithm of marginal likelihood across 5 quantiles (10th, 25th, 50th, 75th, and 90th) from the GFREQ, FREQ, and REQ models for Arizona, California, Illinois, and the entire US.

	10TH QTL	25TH QTL	50TH QTL	75TH QTL	90TH QTL
Arizona−GFREQ	2.22×10^3	2.16×10^3	2.11×10^3	2.18×10^3	2.24×10^3
Arizona−FREQ	1.78×10^3	1.89×10^3	1.94×10^3	1.91×10^2	1.80×10^3
Arizona−REQ	1.50×10^3	1.82×10^3	1.94×10^3	1.85×10^3	1.52×10^3
California−GFREQ	12.3×10^3	11.4×10^3	10.9×10^3	11.2×10^3	11.9×10^3
California−FREQ	9.38×10^3	9.58×10^3	9.61×10^3	9.40×10^3	9.02×10^3
California−REQ	8.92×10^3	9.51×10^3	9.58×10^3	9.06×10^3	8.33×10^3
Illinois−GFREQ	4.90×10^3	4.62×10^3	4.46×10^3	4.64×10^3	4.97×10^3
Illinois−FREQ	3.78×10^3	3.88×10^3	3.90×10^3	3.84×10^3	3.66×10^3
Illinois−REQ	3.41×10^3	3.82×10^3	3.90×10^3	3.70×10^3	3.28×10^3
US−GFREQ	12.4×10^4	11.5×10^4	11.0×10^4	11.4×10^4	12.4×10^4
US−FREQ	9.49×10^4	9.73×10^4	9.79×10^4	9.73×10^4	9.50×10^4
US−REQ	8.83×10^4	9.46×10^4	9.79×10^4	9.47×10^4	8.87×10^4

more flexible approaches in their applied work to improve model fit.

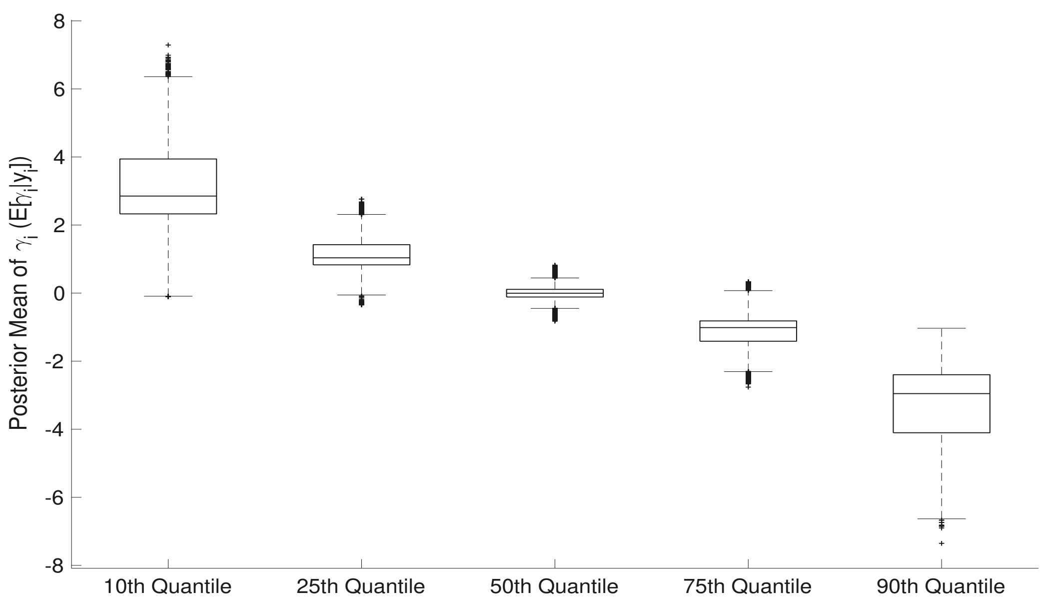

Figure 4: Box plot of posterior mean of γ_i for the US.

Turning attention to the parameter estimates, Table 5 presents the results for the GFREQ model, Table 9 presents the results for the FREQ model, and Table 10 presents the results for

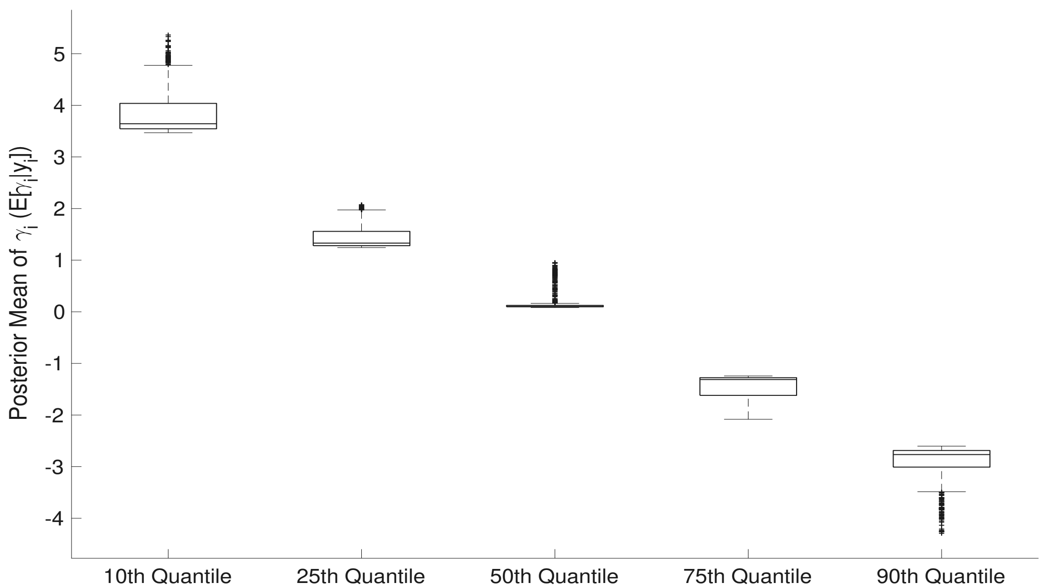

Figure 5: Box plot of posterior mean of γ_i for Illinois.

the REQ model. We find that unemployment is positively associated with residential rental rates. All else equal, a 1 percentage point increase in the fraction of the population that receives unemployment compensation is associated with an increase in monthly residential rental rates of 0.32% (at the 50th quantile). In a 2012 speech to the National Association of Home Builders, Ben Bernanke, then Chairman of the Federal Reserve, stated that "High unemployment and uncertain job prospects may have reduced the willingness of some households to commit to homeownership." By not committing to homeownership, individuals shift their preferences toward renting. The increase in demand for rental units puts upward pressure on prices, explaining the positive result.

What is also notable is the magnitude of the effect, which gets incrementally larger at higher quantiles. Meaning, changes to unemployment are associated with larger changes in rental rates in regions of the U.S. that are more expensive. The effect at the 90th quantile is about 35% larger than that of the 10th quantile. As an economy recovers from a crisis, attention should be paid to the price of rental units. Policymakers may want to focus on limiting the upward pressure on these prices as individuals and families may have a more difficult time recovering from the economic downturn if rents increase.

We also find negative effects from our home mortgage variable (HMrate). Thus, the fraction

Table 5: Results from the GFREQ model for the entire US data–posterior mean (MEAN) and posterior standard deviation (SD) of the parameters.

	10th quantile		25th quantile		50th quantile		75th quantile		90th quantile	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
Intercept	5.64	0.02	5.66	0.02	5.70	0.02	5.75	0.02	5.83	0.02
SSBfrac	-0.83	0.02	-0.88	0.02	-0.89	0.02	-0.91	0.02	-0.91	0.02
FarmFrac	-0.38	0.05	-0.38	0.05	-0.38	0.05	-0.38	0.05	-0.40	0.05
REfrac	0.49	0.03	0.48	0.03	0.46	0.03	0.44	0.03	0.44	0.03
HMrate	-0.22	0.03	-0.23	0.03	-0.23	0.03	-0.22	0.03	-0.23	0.03
AltMinRate	1.98	0.04	1.96	0.04	1.96	0.04	1.96	0.04	1.98	0.04
EnergyRate	0.24	0.02	0.30	0.02	0.34	0.02	0.37	0.02	0.38	0.02
EITCrate	-0.61	0.02	-0.63	0.02	-0.64	0.02	-0.65	0.02	-0.67	0.02
UnempRate	0.26	0.01	0.30	0.01	0.32	0.01	0.33	0.01	0.34	0.01
lAvgAGI	0.20	0.00	0.22	0.00	0.23	0.00	0.23	0.00	0.23	0.00
lreturn	0.08	0.00	0.07	0.00	0.07	0.00	0.07	0.00	0.06	0.00
y11	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00
y12	0.03	0.00	0.03	0.00	0.03	0.00	0.02	0.00	0.01	0.00
y13	0.08	0.00	0.07	0.00	0.07	0.00	0.06	0.00	0.05	0.00
y14	0.10	0.00	0.10	0.00	0.09	0.00	0.08	0.00	0.08	0.00
y15	0.12	0.00	0.11	0.00	0.11	0.00	0.11	0.00	0.10	0.00
y16	0.12	0.00	0.12	0.00	0.12	0.00	0.12	0.00	0.12	0.00
σ	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
φ^2	0.05	0.00	0.04	0.00	0.04	0.00	0.05	0.00	0.05	0.00

of the population taking home mortgage tax deductions is negatively associated with rental prices. And, this result is quite stable across the quantiles. The ability to deduct mortgage interest on individual income taxes makes homeownership more attractive than renting. This result is important in light of the Tax Cuts and Jobs Act (TCJA), which was signed into law in the United States in 2017. The Act lowered the mortgage deduction limit and put a limit on how much an individual can subtract from their taxable income. Our model results suggest that this decrease in home mortgage deductions puts upward pressure on rental prices – a costly unintended consequence. Our results are in line with Hembre and Dantas (2022), who find that reductions in homeownership subsidies increase rental payments.

The other results in Tables 9 and 10 largely align with intuition. Specifically, we find that income is positively associated with rental rates. Average adjusted gross income has a positive effect across the quantiles and the fraction of the population paying alternative minimums (i.e., high-income taxpayers) is also positively associated with rental rates. Whereas, the fraction of the population claiming earned income tax credits (EITC), which represents low-income working individuals, is negatively associated with rental rates. Additionally, the year indicators, which are relative to 2010, are positive and get incrementally larger, capturing aggregate increases in prices. Across all quantiles, relative to 2010, the year 2016 is associated with predicted increases in rental rates of 12%, holding other variables fixed.

4.3.1 Additional Considerations

In this section, we present the GFREQ, FREQ, and REQ results when the sample is restricted to zip codes in Illinois. The model specifications remain the same as before. We chose Illinois (IL) for two reasons: (1) we wish to explore empirical parameter estimates in a smaller sample setting and (2) IL provides extensive variation in land value from expensive metropolitan regions (e.g., Chicago) to rural, farming areas. Figure 6 presents a heat map of zip codes in Illinois and their median rental rates in 2016. There is notable price heterogeneity as distance increases from the Chicago region, where zip codes with prices in the upper decile neighbor zip codes in the 3rd decile. Table 6 presents the GFREQ results, Table 7 presents the FREQ results, and Table 8 presents the REQ results. Importantly, recall from Table 4 that the data support the GFREQ model over the others at all quantiles.

Upon examining the results for the 90th quantile, a large discrepancy between the GFREQ and REQ models becomes apparent. The REQ results shows a posterior mean of -0.23 on the EITCrate variable, whereas the magnitude is -0.48 in the GFREQ model, more than double the size. The REQ model suggests that EITC has the smallest influence on rental rates in expensive areas, but the GFREQ model suggests that expensive rental areas are influenced by EITC just as much as, if not more than, other areas. Another striking difference is the influence of HMrate across quantiles and models. In GFREQ, the posterior mean of HMrate across quantiles is near -1. In particular, it is -1.01 at the 10th quantile. In REQ, on the other hand, the posterior means for HMrate across the quantiles vary quite dramatically, with the lowest value -0.60 appearing at the 10th quantile. The marginal likelihood results in Table 4, however, show overwhelming support for the GFREQ model and the flexibility provided by the GAL distribution and skewness heterogeneity. Researchers or policymakers would have erroneously concluded that the EITC variable has a weaker influence on rental rates in high income areas, when in fact, its effect is quite similar. Additionally, it would be misleading to conclude that home mortgage deduction policies have a smaller effect on

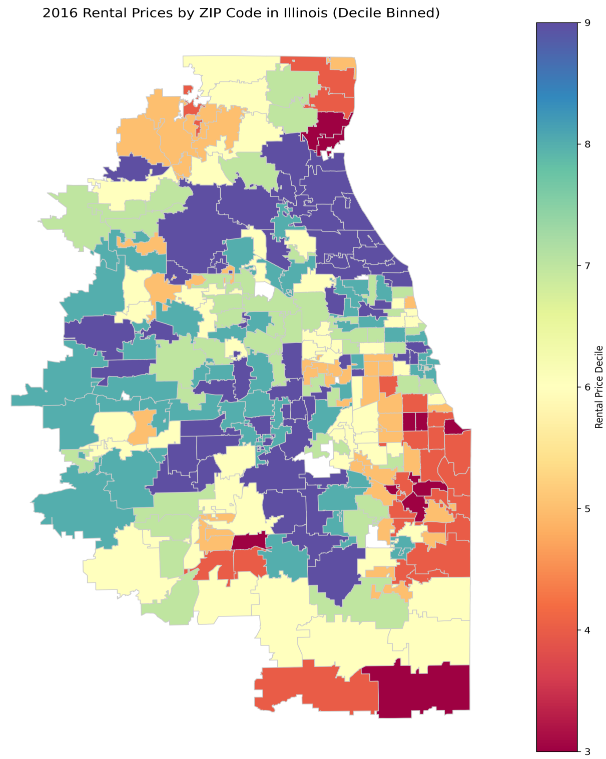

Figure 6: Heat map of rental rates in Illinois in 2016 (binned by decile). The figure is zoomed in to about 70 miles outside of Chicago for clarity.

prices in low-income areas because GFREQ shows it is quite consistent across regions. Overlooking skewness in the data and the advantages of more flexible modeling—especially at the distributional extremes where such benefits are most pronounced—can lead to nontrivial difference in empirical results.

Table 6: Results from the GFREQ model for Illinois–posterior mean (MEAN) and posterior standard deviation (SD) of the parameters.

	10th quantile		25th quantile		50th quantile		75th quantile		90th quantile	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
Intercept	5.66	0.15	5.58	0.14	5.60	0.14	5.58	0.14	5.49	0.12
SSBfrac	−1.55	0.16	−1.68	0.14	−1.67	0.15	−1.72	0.14	−1.76	0.13
FarmFrac	0.95	0.28	1.24	0.28	1.31	0.30	1.30	0.26	1.37	0.23
REfrac	0.93	0.22	1.06	0.20	1.23	0.22	1.03	0.21	0.86	0.16
HMrate	−1.01	0.23	−1.18	0.21	−1.15	0.24	−1.02	0.21	−0.92	0.17
AltMinRate	1.18	0.23	0.94	0.21	0.80	0.23	0.69	0.22	0.49	0.19
EnergyRate	1.12	0.15	0.99	0.13	0.98	0.15	0.78	0.13	0.65	0.11
EITCrate	−0.51	0.12	−0.49	0.11	−0.27	0.11	−0.42	0.11	−0.48	0.09
UnempRate	−0.17	0.12	−0.10	0.11	−0.02	0.12	−0.02	0.11	−0.01	0.08
lAvgAGI	0.24	0.03	0.27	0.02	0.27	0.03	0.30	0.03	0.32	0.02
lreturn	0.08	0.01	0.09	0.01	0.07	0.01	0.08	0.01	0.09	0.01
y11	0.00	0.01	−0.01	0.01	−0.01	0.01	−0.02	0.01	−0.03	0.01
y12	0.02	0.01	0.01	0.01	0.01	0.01	−0.01	0.01	−0.02	0.01
y13	0.04	0.01	0.03	0.01	0.04	0.01	0.02	0.01	0.00	0.01
y14	0.04	0.01	0.03	0.01	0.04	0.01	0.02	0.01	0.00	0.01
y15	0.04	0.01	0.03	0.01	0.05	0.01	0.03	0.01	0.00	0.01
y16	0.03	0.01	0.02	0.01	0.05	0.01	0.03	0.01	0.01	0.01
σ	0.01	0.00	0.01	0.00	0.02	0.00	0.01	0.00	0.01	0.00
φ^2	0.04	0.00	0.04	0.00	0.04	0.00	0.04	0.00	0.04	0.00

5 Conclusion

[POINTS TO CONSIDER: the conclusion needs strengthening—the first three paragraphs is too brief and underdeveloped compared to the abstract and introduction.]

This article has considered the Bayesian analysis of a random effects quantile regression model for panel data under the generalized asymmetric Laplace distribution, which relaxes the dependence of distributional skewness on the quantile parameter. New computationally efficient MCMC simulation algorithms have been developed and employed for parameter estimation and model comparison in the GFREQ, FREQ, and REQ versions of the model. Key to the improved properties of our simulation algorithm is the idea of carefully designed parameter blocking. Various features of the proposed modeling framework and estimation methodology have been studied in simulation studies.

The paper has also devoted considerable attention to studying the behavior of U.S. residential rental rates following the Global Financial Crisis. Our methodology fits this purpose very well due

Table 7: Results from the FREQ model for Illinois–posterior mean (MEAN) and posterior standard deviation (SD) of the parameters.

	10th quantile		25th quantile		50th quantile		75th quantile		90th quantile	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
Intercept	5.49	0.12	5.45	0.13	5.43	0.13	5.49	0.13	5.63	0.13
SSBfrac	−1.41	0.13	−1.55	0.13	−1.66	0.13	−1.78	0.13	−1.77	0.13
FarmFrac	1.03	0.24	1.25	0.26	1.41	0.26	1.37	0.25	1.13	0.26
REfrac	0.65	0.17	0.77	0.17	0.87	0.18	0.97	0.18	0.97	0.18
HMrate	−0.81	0.18	−0.91	0.18	−1.01	0.19	−1.06	0.19	−0.99	0.19
AltMinRate	0.98	0.19	0.81	0.19	0.66	0.21	0.59	0.21	0.64	0.22
EnergyRate	1.12	0.12	1.07	0.12	0.99	0.12	0.82	0.12	0.68	0.13
EITCrate	−0.52	0.10	−0.51	0.10	−0.51	0.10	−0.51	0.10	−0.46	0.10
UnempRate	−0.03	0.09	−0.03	0.09	−0.01	0.09	0.03	0.10	−0.02	0.10
lAvgAGI	0.27	0.02	0.28	0.02	0.30	0.02	0.31	0.02	0.31	0.03
lreturn	0.09	0.01	0.09	0.01	0.09	0.01	0.09	0.01	0.08	0.01
y11	0.01	0.01	0.00	0.01	−0.01	0.01	−0.02	0.01	−0.03	0.01
y12	0.02	0.01	0.02	0.01	0.01	0.01	−0.01	0.01	−0.02	0.01
y13	0.05	0.01	0.04	0.01	0.03	0.01	0.01	0.01	−0.01	0.01
y14	0.04	0.01	0.04	0.01	0.03	0.01	0.01	0.01	−0.01	0.01
y15	0.04	0.01	0.03	0.01	0.03	0.01	0.02	0.01	0.00	0.01
y16	0.03	0.01	0.03	0.01	0.02	0.01	0.02	0.01	0.00	0.01
σ	0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.00
γ	1.74	0.06	0.55	0.03	−0.10	0.05	−0.81	0.04	−2.01	0.09
φ^2	0.04	0.00	0.04	0.00	0.04	0.00	0.04	0.00	0.04	0.00

to the strong right skew of rental rates and the extensive heterogeneity at the zip-code level across different regions. Our results reveal that unemployment has positive effects on rental rates and mortgage deductions have a negative impact. Regions of the U.S. characterized by high unemployment also exhibited declines in homeownership, leading to an increase in demand for rental units and putting upward pressure on prices. The negative effect of mortgage deductions sheds light on the unintended consequences of the Tax Cuts and Jobs Act (TCJA) as a potential contributor to the large increases in rental prices since 2017.

Based on our model comparisons, we find that the data overwhelmingly support the GFREQ model in various sub-samples and at all quantiles, especially away from the median, suggesting that researchers interested in the tails of the distribution could find the more flexible GAL modeling framework decidedly more useful. And while we hope that the methodology can be effective in a variety of practical applications where heterogeneity and distributional asymmetry matter, in future work we also intend to study possible links and extensions to other approaches including

Table 8: Results from REQ model for Illinois–posterior mean (MEAN) and posterior standard deviation (SD) of the parameters.

	10th quantile		25th quantile		50th quantile		75th quantile		90th quantile	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
Intercept	5.56	0.13	5.53	0.13	5.41	0.13	5.54	0.14	5.79	0.13
SSBfrac	−1.12	0.15	−1.37	0.13	−1.71	0.13	−1.78	0.14	−1.57	0.14
FarmFrac	0.88	0.22	1.09	0.24	1.46	0.26	1.31	0.25	0.90	0.24
REfrac	0.50	0.18	0.61	0.17	0.92	0.18	0.96	0.18	0.92	0.17
HMrate	−0.60	0.19	−0.76	0.18	−1.04	0.19	−0.93	0.19	−0.76	0.18
AltMinRate	1.15	0.20	1.00	0.19	0.58	0.21	0.47	0.22	0.68	0.23
EnergyRate	1.31	0.14	1.22	0.12	0.93	0.12	0.60	0.13	0.36	0.12
EITCrate	−0.49	0.11	−0.51	0.10	−0.51	0.10	−0.43	0.10	−0.23	0.10
UnempRate	−0.17	0.10	−0.11	0.09	0.00	0.09	−0.06	0.10	−0.12	0.09
lAvgAGI	0.22	0.02	0.25	0.02	0.31	0.02	0.32	0.03	0.30	0.03
lreturn	0.10	0.01	0.09	0.01	0.09	0.01	0.08	0.01	0.06	0.01
y11	0.02	0.01	0.01	0.01	−0.01	0.01	−0.04	0.01	−0.06	0.01
y12	0.04	0.01	0.03	0.01	0.00	0.01	−0.03	0.01	−0.05	0.01
y13	0.07	0.01	0.05	0.01	0.02	0.01	−0.01	0.01	−0.03	0.01
y14	0.07	0.01	0.05	0.01	0.02	0.01	−0.01	0.01	−0.03	0.01
y15	0.06	0.01	0.04	0.01	0.02	0.01	0.00	0.01	−0.02	0.01
y16	0.04	0.01	0.03	0.01	0.02	0.01	0.01	0.01	−0.01	0.01
σ	0.01	0.00	0.02	0.00	0.02	0.00	0.02	0.00	0.01	0.00
φ^2	0.05	0.00	0.04	0.00	0.04	0.00	0.04	0.00	0.04	0.00

tilting, skew-distribution modeling, and expectiles, among others.

To conclude the paper, we outline two promising directions for future research. First, motivated by our empirical application, it appears natural to extend the proposed FREQ model to incorporate spatial dependence. In the spirit of spatial Durbin model (Anselin, 1988) and its extension to panel data settings (Beer and Riedl, 2012), the spatial-FREQ (S-FREQ) model can be written as,

$$y_{it} = \rho \sum_{j=1}^{n} w_{ij} y_{jt} + x'_{it}\beta + \sum_{j=1}^{n} w_{ij}(x'_{ij}\zeta) + z'_{it}\alpha_i + \varepsilon_{it}, \tag{16}$$

where, as compared to the FREQ model, there are three additional parameters: the spatial autocorrelation coefficient ρ, the spatial weight elements w_{ij} of the spatial weight matrix W, and the parameter vector ζ. As depicted in equation (16), the spatial weights enter the model through dependent variable and covariates, while the error term retains the same distributional assumptions as the original model , i.e., $\varepsilon_{it} \overset{iid}{\sim} \text{GAL}(0, \sigma, p_0, \gamma)$. This permits the continued use of the mixture representation for efficient MCMC estimation. The S-FREQ model offers considerable flexibility, with various nested specifications obtainable through parameter restrictions. For instance, when

$\rho = 0$, the model simplifies to a FREQ model that includes a spatial lag on the regressors. This is the simplest case because the model can be estimated using Algorithm 1 with an extra component arising from the spatial lag on the regressors. Conversely, $\gamma = 0$ results in a spatial quantile panel autoregressive model, as discussed in Yu et al. (2025).

Second, we have employed the marginal likelihood to demonstrate that both the generalized FREQ and FREQ models offer superior model fit compared to the REQ model. While a model with a higher marginal likelihood typically indicates better overall out-of-sample predictive performance, it is essential to formally assess whether this superiority holds at specific conditional quantiles of interest. To address this, future research could involve conducting an out-of-sample prediction exercise using a hold-out sample. The out-of-sample quantile loss could then be computed using an appropriate scoring rule, such as the one proposed by Gneiting and Raftery (2007), which has been widely adopted and well-regarded in the literature for its robustness in quantile prediction evaluations.

S.K.: Derivation of conditional density for FREQ can be removed, I have kept it here so that cross referencing does not break down.

Appendix A Conditional Densities in the FREQ model

The joint posterior density of the FREQ model of Section 2.1 is given by

$$
\begin{aligned}
&\pi(\Theta|y) \propto f(y|\Theta) \times \pi(\alpha|\Omega) \times \pi(\nu) \times \pi(h) \times \pi(\beta) \times \pi(\sigma) \times \pi(\Omega) \times \pi(\gamma) \\
&\propto \prod_{i=1}^{n} \Bigg[|\Lambda_i|^{-1/2} \exp\bigg\{ -\frac{1}{2}\Big(y_i - X_i\beta - Z_i\alpha_i - A\nu_i - C|\gamma|h_i\Big)' \Lambda_i^{-1} \Big(y_i - X_i\beta - Z_i\alpha_i - A\nu_i \\
&\quad - C|\gamma|h_i\Big)\bigg\} \times |\Omega|^{-1/2} \exp\bigg\{ -\frac{1}{2}\, \alpha_i'\Omega^{-1}\alpha_i \bigg\} \times \bigg(\sigma^{-T_i} \exp\bigg\{ -\sum_{j=1}^{T_i} \frac{\nu_{it}}{\sigma} \bigg\}\bigg) \\
&\quad \times \bigg(\sigma^{-T_i} \exp\bigg\{ -\frac{h_i'h_i}{2\sigma^2} \bigg\}\bigg)\Bigg] \times \exp\bigg\{ -\frac{1}{2}(\beta-\beta_0)' B_0^{-1} (\beta - \beta_0) \bigg\} \\
&\quad \times \bigg(\sigma^{-\frac{n_0}{2}-1} \exp\bigg\{ -\frac{d_0}{2\sigma} \bigg\}\bigg) \times |\Omega|^{-(r_0+l+1)/2} \exp\bigg\{ -\frac{1}{2}\,\mathrm{tr}(\Omega^{-1}R_0) \bigg\}, \qquad (17)
\end{aligned}
$$

from which we derive the conditional densities in the same order as the steps in Algorithm 1 with parameter dependence on the quantile being suppressed for notational simplicity.

(1) The parameters (β, α) are sampled in one block to account for correlation between them and reduce autocorrelations in the Markov chain (see Chib and Carlin, 1999). The joint conditional density of (β, α) can be expressed as

$$
\begin{aligned}
\pi(\beta, \alpha|y, \nu, h, \sigma, \gamma, \Omega) &= \pi(\beta|y, \nu, h, \sigma, \gamma, \Omega)\pi(\alpha|y, \beta, \nu, h, \sigma, \gamma, \Omega) \\
&= \pi(\beta|y, \nu, h, \sigma, \gamma, \Omega) \prod_{i=1}^{n} \pi(\alpha_i|y, \beta, \nu, h, \sigma, \gamma, \Omega).
\end{aligned}
$$

We first sample β marginally of α and then α conditionally on β and the other parameters.

(a) To find the conditional density $\pi(\beta|y, \nu, h, \sigma, \gamma, \Omega)$, we integrate out (α_i, u_i) from the model

$$
y_i = X_i\beta + Z_i\alpha_i + A\nu_i + C|\gamma|h_i + \Lambda_i^{1/2}u_i,
$$

where $\alpha_i \sim N(0_l, \Omega_{l\times l})$ and $u_i \sim N(0_{T_i}, I_{T_i})$, implying that $y_i|\beta, \nu, h, \sigma, \gamma, \Omega \sim N(X_i\beta + A\nu_i + C|\gamma|h_i,\ Z_i\Omega Z_i' + \Lambda_i)$ for $i = 1, \ldots, n$. Consequently, we have that $\beta|y, \nu, h, \sigma, \gamma, \Omega \sim N(\tilde{\beta}, \tilde{B})$ with

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

(b) The full conditional density $\pi(\alpha_i|y, \beta, \nu, h, \sigma, \gamma, \Omega) \propto f(y_i|\beta, \alpha_i, \nu, h, \sigma, \gamma, \Omega)\pi(\alpha_i|\Omega)$ can then be seen to be $\alpha_i|y, \beta, \nu, h, \sigma, \gamma, \Omega \sim N(\tilde{\alpha}_i, \tilde{A}_i)$, for $i = 1, \ldots, n$, with

$$\tilde{A}_i = \left(Z_i'\Lambda_i^{-1}Z_i + \Omega^{-1}\right)^{-1} \quad \text{and} \quad \tilde{\alpha}_i = \tilde{A}_i \left(Z_i'\Lambda_i^{-1}(y_i - X_i\beta - A\nu_i - C|\gamma|h_i)\right).$$

(2) The parameters (σ, γ) are sampled as a single block, marginally of (ν, h), from a conditional density that is proportional to the product of Equation (3) and the priors in Equation (6). Because that density is not tractable, simulation requires a Metropolis-Hastings step. While joint sampling of (σ, γ) reduces the inefficiency factors considerably, an increase in the overall algorithmic efficiency also depends on moderating the computational costs of block sampling. We achieve this by employing a random walk proposal $(\sigma^{\dagger}, \gamma^{\dagger}) \sim TN_{(0,\infty)\times(L,U)}\big((\sigma_c, \gamma_c), \iota^2\hat{D}\big)$, where the limits of truncation are as in Equation (2), (σ_c, γ_c) are the current values, ι is a tuning factor and $\hat{D}$ is obtained at the start of the Markov chain as the modal dispersion matrix of (σ, γ) from the GAL log-likelihood of an approximate pooled data model with β set to the pooled ordinary least squares estimate. This construction is far faster and less computationally intensive than a tailored independence chain that involves updating the mean and covariance matrix at every MCMC step, yet it yields good mixing and low inefficiencies as shown in the Section 3. The algorithm of Botev (2016) is employed to simulate efficiently from the bivariate truncated normal proposal. The proposed $(\sigma^{\dagger}, \gamma^{\dagger})$ is accepted as the next MCMC draw with probability

$$\alpha_{MH}(\sigma_c, \gamma_c; \sigma^{\dagger}, \gamma^{\dagger}|y, \cdot) = \min\left\{1, \left[\frac{f(y, \alpha|\beta, \sigma^{\dagger}, \gamma^{\dagger})\, \pi(\beta, \sigma^{\dagger}, \gamma^{\dagger})}{f(y, \alpha|\beta, \sigma_c, \gamma_c)\, \pi(\beta, \alpha, \sigma_c, \gamma_c)}\; \frac{q(\sigma_c, \gamma_c|(\sigma^{\dagger}, \gamma^{\dagger}), \iota^2\hat{D})}{q(\sigma^{\dagger}, \gamma^{\dagger}|(\sigma_c, \gamma_c), \iota^2\hat{D})}\right]\right\},$$

where, $f(\cdot)$ is the complete data likelihood given by Equation (3), $\pi(\cdot)$ denotes the densities of the priors in Equation (6), and $q(\cdot)$ represents the density of bivariate truncated normal proposal. If $(\sigma^{\dagger}, \gamma^{\dagger})$ is not accepted, the current value (σ_c, γ_c) is repeated as the next MCMC draw.

(3) To derive the conditional posterior density of ν_{it}, we rely on the GAL mixture representation and work element-wise to obtain

$$\begin{aligned}
&\pi(\nu_{it}|y_{it}, \beta, \alpha_i, h_{it}, \sigma, \gamma) \\
&\propto \nu_{it}^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}\left[\frac{(y_{it} - x_{it}'\beta - z_{it}'\alpha_i - A\nu_{it} - C|\gamma|h_{it})^2}{\sigma B \nu_{it}}\right] - \frac{\nu_{it}}{\sigma}\right\} \\
&\propto \nu_{it}^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}\left[\frac{(y_{it} - x_{it}'\beta - z_{it}'\alpha_i - C|\gamma|h_{it})^2}{\sigma B}\, \nu_{it}^{-1} + \left(\frac{A^2}{\sigma B} + \frac{2}{\sigma}\right)\nu_{it}\right]\right\} \\
&\propto \nu_{it}^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}\left[\chi \nu_{it}^{-1} + \psi_{\nu_{it}}\nu_{it}\right]\right\},
\end{aligned}$$

where we have used the notation

$$\chi = \left(\frac{A^2}{\sigma B} + \frac{2}{\sigma}\right) \qquad \text{and} \qquad \psi_{\nu_{it}} = \frac{(y_{it} - x_{it}'\beta - s_{it}'\alpha_i - C|\gamma|h_{it})^2}{\sigma B}.$$

Therefore, $\nu_{it}|y_{it},\beta,\alpha_i,h_{it},\sigma,\gamma \sim GIG(\frac{1}{2},\chi,\psi_{\nu_{it}})$ for all i and t.

(4) The conditional density of h_{it} is derived element-wise from the GAL mixture representation:

$$\begin{aligned}
&\pi(h_{it}|y_{it},\beta,\nu_{it},\sigma,\gamma) \\
&\propto \exp\left\{-\frac{1}{2}\left[\frac{(y_{it} - x_{it}'\beta - z_{it}\alpha_i - A\nu_{it} - C|\gamma|h_{it})^2}{\sigma B\nu_{it}} + \frac{h_{it}^2}{\sigma^2}\right]\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left[\left(\frac{1}{\sigma^2} + \frac{C^2\gamma^2}{\sigma B\nu_{it}}\right)h_{it}^2 - \frac{2C|\gamma|(y_i - x_{it}'\beta - z_{it}'\alpha_i - A\nu_{it})}{\sigma B\nu_{it}}h_{it}\right]\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left[(\sigma_{h_{it}}^2)^{-1}h_{it}^2 - 2(\sigma_{h_{it}}^2)^{-1}\mu_{h_{it}}h_{it}\right]\right\} \\
&\propto \exp\left\{-\frac{1}{2}(\sigma_{h_{it}}^2)^{-1}(h_{it} - \mu_{h_{it}})^2\right\}
\end{aligned}$$

with

$$\sigma_{h_{it}}^2 = \left(\frac{1}{\sigma^2} + \frac{C^2\gamma^2}{\sigma B\nu_{it}}\right)^{-1} \qquad \text{and} \qquad \mu_{h_{it}} = \sigma_{h_{it}}^2\left(\frac{C|\gamma|(y_{it} - x_{it}'\beta - z_{it}'\alpha_i - A\nu_{it})}{\sigma B\nu_{it}}\right).$$

We thus have the half-normal kernel, whereby $h_{it}|y_{it},\beta,\nu_{it},\sigma,\gamma \sim N^+(\mu_{h_{it}},\sigma_{h_{it}}^2)$ for all i and t.

(5) Derivation of the full conditional distribution of Ω is straightforward since

$$\begin{aligned}
\pi(\Omega|y,\alpha) &\propto \prod_{i=1}^{n}\Big[\pi(\alpha_i|\Omega)\Big] \times \pi(\Omega) \\
&\propto \prod_{i=1}^{n}\left[|\Omega|^{-1/2}\exp\left\{-\frac{1}{2}\,\alpha_i'\Omega^{-1}\alpha_i\right\}\right] \times |\Omega|^{-(r_0+l+1)/2}\exp\left\{-\frac{1}{2}\,\mathrm{tr}(\Omega^{-1}R_0)\right\} \\
&\propto |\Omega|^{-(n+r_0+l+1)/2}\exp\left\{-\frac{1}{2}\,\mathrm{tr}\left[\Omega^{-1}\left(\sum_{i=1}^{n}\alpha_i\alpha_i' + R_0\right)\right]\right\},
\end{aligned}$$

whereby $\Omega|y,\alpha \sim IW(\tilde{\omega},\tilde{R})$ with $\tilde{\omega} = n + r_0$ and $\tilde{R} = \sum_{i=1}^{n}\alpha_i\alpha_i' + R_0$.

Appendix B Conditional Densities in the Generalized FREQ model

The joint posterior density of the Gen-FREQ model of Section 2.1 is given by

$$
\begin{aligned}
&\pi(\Theta|y) \propto f(y|\Theta) \times \pi(\alpha|\Omega) \times \pi(\nu) \times \pi(h) \times \pi(\beta) \times \pi(\sigma) \times \pi(\Omega) \times \pi(\gamma) \\
&\propto \prod_{i=1}^{n} \Bigg[|\Lambda_i|^{-1/2} \exp\bigg\{ -\frac{1}{2}\Big(y_i - X_i\beta - Z_i\alpha_i - A\nu_i - C|\gamma_i|h_i\Big)' \Lambda_i^{-1} \Big(y_i - X_i\beta - Z_i\alpha_i - A\nu_i \\
&\quad - C|\gamma_i|h_i\Big)\bigg\} \times |\Omega|^{-1/2} \exp\bigg\{ -\frac{1}{2}\,\alpha_i'\Omega^{-1}\alpha_i \bigg\} \times \bigg(\sigma^{-T_i} \exp\bigg\{ -\sum_{j=1}^{T_i} \frac{\nu_{it}}{\sigma} \bigg\} \bigg) \\
&\quad \times \bigg(\sigma^{-T_i} \exp\bigg\{ -\frac{h_i'h_i}{2\sigma^2} \bigg\} \bigg) \Bigg] \times \exp\bigg\{ -\frac{1}{2}(\beta-\beta_0)' B_0^{-1} (\beta-\beta_0) \bigg\} \\
&\quad \times \bigg(\sigma^{-\frac{n_0}{2}-1} \exp\bigg\{ -\frac{d_0}{2\sigma} \bigg\} \bigg) \times |\Omega|^{-(r_0+l+1)/2} \exp\bigg\{ -\frac{1}{2}\,\mathrm{tr}(\Omega^{-1}R_0) \bigg\}, \qquad (18)
\end{aligned}
$$

from which we derive the conditional densities in the same order as the steps in Algorithm 1 with parameter dependence on the quantile being suppressed for notational simplicity.

(1) The parameters (β, α) are sampled in one block to account for correlation between them and reduce autocorrelations in the Markov chain (see Chib and Carlin, 1999). The joint conditional density of (β, α) can be expressed as

$$
\begin{aligned}
\pi(\beta, \alpha|y, \nu, h, \sigma, \gamma, \Omega) &= \pi(\beta|y, \nu, h, \sigma, \gamma, \Omega)\pi(\alpha|y, \beta, \nu, h, \sigma, \gamma, \Omega) \\
&= \pi(\beta|y, \nu, h, \sigma, \gamma, \Omega) \prod_{i=1}^{n} \pi(\alpha_i|y, \beta, \nu, h, \sigma, \gamma_i, \Omega).
\end{aligned}
$$

We first sample β marginally of α and then α conditionally on β and the other parameters.

(a) To find the conditional density $\pi(\beta|y, \nu, h, \sigma, \gamma, \Omega)$, we integrate out (α_i, u_i) from the model

$$
y_i = X_i\beta + Z_i\alpha_i + A\nu_i + C|\gamma_i|h_i + \Lambda_i^{1/2}u_i,
$$

where $\alpha_i \sim N(0_l, \Omega_{l\times l})$ and $u_i \sim N(0_{T_i}, I_{T_i})$, implying that $y_i|\beta, \nu, h, \sigma, \gamma_i, \Omega \sim N(X_i\beta + A\nu_i + C|\gamma_i|h_i,\ Z_i\Omega Z_i' + \Lambda_i)$ for $i = 1, \ldots, n$. Consequently, we have that $\beta|y, \nu, h, \sigma, \gamma, \Omega \sim N(\tilde{\beta}, \tilde{B})$ with

$$
\tilde{B} = \left(\sum_{i=1}^{n} X_i' V_i^{-1} X_i + B_0^{-1} \right)^{-1} \quad \text{and} \quad \tilde{\beta} = \tilde{B} \left(\sum_{i=1}^{n} X_i' V_i^{-1} (y_i - A\nu_i - C|\gamma_i|h_i) + B_0^{-1}\beta_0 \right).
$$

(b) The full conditional density $\pi(\alpha_i|y, \beta, \nu, h, \sigma, \gamma_i, \Omega) \propto f(y_i|\beta, \alpha_i, \nu, h, \sigma, \gamma_i, \Omega)\pi(\alpha_i|\Omega)$ can then be seen to be $\alpha_i|y, \beta, \nu, h, \sigma, \gamma_i, \Omega \sim N(\tilde{\alpha}_i, \tilde{A}_i)$, for $i = 1, \ldots, n$, with

$$
\tilde{A}_i = \left(Z_i' \Lambda_i^{-1} Z_i + \Omega^{-1} \right)^{-1} \quad \text{and} \quad \tilde{\alpha}_i = \tilde{A}_i \left(Z_i' \Lambda_i^{-1} (y_i - X_i\beta - A\nu_i - C|\gamma_i|h_i) \right).
$$

(2) The parameters (σ, γ) are sampled in a single block by first sampling σ marginally of (γ, ν, h) and then sampling γ_i marginally of (ν, h) and conditionally on all the other parameters. The conditional density of σ is proportional to the product of integrated likelihood in equation (9) and priors in (6). Since the resultant conditional density is intractable, a Metropolis-Hastings step is required for simulation. We employ a random walk proposal $\sigma^\dagger \sim TN_{(0,\infty)}\big(\sigma_c, \iota_\sigma^2 \hat{D}_\sigma\big)$, σ_c is the current draw, ι is a tuning factor and $\hat{D}_\sigma$ is obtained at the start of the Markov chain as the modal dispersion of σ from the integrated log-likelihood of an approximate pooled data model with β set to the pooled ordinary least squares estimate.[9] This construction is far faster and less computationally intensive than a tailored independence chain that involves updating the mean and covariance matrix at every MCMC step, yet it yields good mixing and low inefficiencies as shown in the Section 3. The algorithm of Botev (2016) is employed to simulate efficiently from the truncated normal proposal. The proposed $\sigma^\dagger$ accepted as the next MCMC draw with probability

$$\alpha_{MH}(\sigma_c; \sigma^\dagger | y, \cdot) = \min\left\{1, \left[\frac{\prod_{i=1}^{n} \int_L^U f(y_i, \alpha_i | \beta, \sigma^\dagger, \gamma_i) d\gamma_i}{\prod_{i=1}^{n} \int_L^U f(y_i, \alpha_i | \beta, \sigma, \gamma_i) d\gamma_i} \frac{\pi(\sigma^\dagger)}{\pi(\sigma_c)} \frac{q(\sigma^c | \sigma^\dagger, \iota_\sigma^2 \hat{D}_\sigma)}{q(\sigma^\dagger | \sigma_c, \iota_\sigma^2 \hat{D}_\sigma)}\right]\right\}$$

The expressions $\int_L^U f(y_i, \alpha_i | \beta, \sigma^\dagger, \gamma_i) d\gamma_i$ is obtained by numerical integration using adaptive quadrature method (see Shampine, 2008). As noted earlier, the individual contribution to the integrated likelihood are independent across the clusters, hence integration can be parallelized across multiple threads to significantly reduce computational time.

The shape parameter γ_i is sampled marginally of (ν_i, h_i) from the conditional density which is proportional to the likelihood $f_{GAL}(y|\beta, \alpha, \sigma, \gamma_i)$ times the prior distributions $\pi(\beta, \alpha, \sigma, \gamma_i)$ given by equation (3) and equation (6),respectively. Collecting terms involving γ does not yield a tractable distribution, so γ_i is also sampled using a random-walk MH algorithm. The proposed draw $\gamma_i^\dagger$ is generated from truncated normal distribution $TN_{(L,U)}\big(\gamma_{i,c}, \iota_\gamma^2 \hat{D}_\gamma\big)$, $\gamma_{i,c}$ is the current draw, ι_γ is a tuning factor and $\hat{D}_\gamma$ is obtained at the start of the Markov chain as the modal dispersion of the likelihood $f(y_i, \alpha_i | \beta, \sigma)$. The proposed $\gamma_i^\dagger$ is accepted as the next MCMC draw with probability

$$\alpha_{MH}(\gamma_i^\dagger, \gamma_{i,c} | y_i, .) = \min\left\{1, \left[\frac{f(y_i, \alpha_i | \beta, \sigma, \gamma_i^\dagger)}{f(y_i, \alpha_i | \beta, \sigma, \gamma_{i,c})} \frac{\pi(\gamma_i^\dagger)}{\pi(\gamma_{i,c})} \frac{q(\gamma_{i,c} | \gamma_i^\dagger, \iota_\gamma^2 \hat{D}_\gamma)}{q(\gamma_i^\dagger | \gamma_{i,c}, \iota_\gamma^2 \hat{D}_\gamma)}\right]\right\}$$

where, $f(\cdot)$ is the complete data likelihood given by Equation (3), $\pi(\cdot)$ denotes the densities of the priors in Equation (6), and $q(\cdot)$ represents the density of truncated normal proposal. If $\gamma_i^\dagger$ is not

[9]In practice, one can also set the value of β to first-stage posterior mean obtained from pre-sample training.

accepted, the current value $\gamma_{i,c}$ is repeated as the next MCMC draw.

(3) To derive the conditional posterior density of ν_{it}, we rely on the GAL mixture representation and work element-wise to obtain

$$
\begin{aligned}
&\pi(\nu_{it}|y_{it},\beta,\alpha_i,h_{it},\sigma,\gamma_i)\\
&\propto \nu_{it}^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left[\frac{(y_{it}-x_{it}'\beta-z_{it}'\alpha_i-A\nu_{it}-C|\gamma_i|h_{it})^2}{\sigma B\nu_{it}}\right]-\frac{\nu_{it}}{\sigma}\right\}\\
&\propto \nu_{it}^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left[\frac{(y_{it}-x_{it}'\beta-z_{it}'\alpha_i-C|\gamma_i|h_{it})^2}{\sigma B}\,\nu_{it}^{-1}+\left(\frac{A^2}{\sigma B}+\frac{2}{\sigma}\right)\nu_{it}\right]\right\}\\
&\propto \nu_{it}^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left[\chi\nu_{it}^{-1}+\psi_{\nu_{it}}\nu_{it}\right]\right\},
\end{aligned}
$$

where we have used the notation

$$
\chi=\left(\frac{A^2}{\sigma B}+\frac{2}{\sigma}\right)\qquad\text{and}\qquad\psi_{\nu_{it}}=\frac{(y_{it}-x_{it}'\beta-s_{it}'\alpha_i-C|\gamma_i|h_{it})^2}{\sigma B}.
$$

Therefore, $\nu_{it}|y_{it},\beta,\alpha_i,h_{it},\sigma,\gamma_i\sim GIG(\frac{1}{2},\chi,\psi_{\nu_{it}})$ for all i and t.

(4) The conditional density of h_{it} is derived element-wise from the GAL mixture representation:

$$
\begin{aligned}
&\pi(h_{it}|y_{it},\beta,\nu_{it},\sigma,\gamma_i)\\
&\propto \exp\left\{-\frac{1}{2}\left[\frac{(y_{it}-x_{it}'\beta-z_{it}\alpha_i-A\nu_{it}-C|\gamma_i|h_{it})^2}{\sigma B\nu_{it}}+\frac{h_{it}^2}{\sigma^2}\right]\right\}\\
&\propto \exp\left\{-\frac{1}{2}\left[\left(\frac{1}{\sigma^2}+\frac{C^2\gamma_i^2}{\sigma B\nu_{it}}\right)h_{it}^2-\frac{2C|\gamma_i|(y_i-x_{it}'\beta-z_{it}'\alpha_i-A\nu_{it})}{\sigma B\nu_{it}}h_{it}\right]\right\}\\
&\propto \exp\left\{-\frac{1}{2}\left[(\sigma_{h_{it}}^2)^{-1}h_{it}^2-2(\sigma_{h_{it}}^2)^{-1}\mu_{h_{it}}h_{it}\right]\right\}\\
&\propto \exp\left\{-\frac{1}{2}(\sigma_{h_{it}}^2)^{-1}(h_{it}-\mu_{h_{it}})^2\right\}
\end{aligned}
$$

with

$$
\sigma_{h_{it}}^2=\left(\frac{1}{\sigma^2}+\frac{C^2\gamma_i^2}{\sigma B\nu_{it}}\right)^{-1}\qquad\text{and}\qquad\mu_{h_{it}}=\sigma_{h_{it}}^2\left(\frac{C|\gamma_i|(y_{it}-x_{it}'\beta-z_{it}'\alpha_i-A\nu_{it})}{\sigma B\nu_{it}}\right).
$$

We thus have the half-normal kernel, whereby $h_{it}|y_{it},\beta,\nu_{it},\sigma,\gamma_i\sim N^+(\mu_{h_{it}},\sigma_{h_{it}}^2)$ for all i and t.

(5) Derivation of the full conditional distribution of Ω is straightforward since

$$\begin{aligned}
\pi(\Omega|y,\alpha) &\propto \prod_{i=1}^{n}\Big[\pi(\alpha_i|\Omega)\Big] \times \pi(\Omega) \\
&\propto \prod_{i=1}^{n}\left[|\Omega|^{-1/2}\exp\left\{-\frac{1}{2}\,\alpha_i'\Omega^{-1}\alpha_i\right\}\right] \times |\Omega|^{-(r_0+l+1)/2}\exp\left\{-\frac{1}{2}\,\mathrm{tr}(\Omega^{-1}R_0)\right\} \\
&\propto |\Omega|^{-(n+r_0+l+1)/2}\exp\left\{-\frac{1}{2}\,\mathrm{tr}\left[\Omega^{-1}\left(\sum_{i=1}^{n}\alpha_i\alpha_i' + R_0\right)\right]\right\},
\end{aligned}$$

whereby $\Omega|y,\alpha \sim IW(\tilde{\omega},\tilde{R})$ with $\tilde{\omega} = n + r_0$ and $\tilde{R} = \sum_{i=1}^{n}\alpha_i\alpha_i' + R_0$.

Appendix C Figures and Tables

Appendix D Application Appendix

Table 9: Results from the FREQ model for the entire US data–posterior mean (MEAN) and posterior standard deviation (SD) of the parameters.

	10th quantile		25th quantile		50th quantile		75th quantile		90th quantile	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
Intercept	5.63	0.02	5.63	0.02	5.65	0.02	5.70	0.02	5.76	0.02
SSBfrac	−0.89	0.02	−0.94	0.02	−0.95	0.02	−0.97	0.02	−0.97	0.02
FarmFrac	−0.41	0.05	−0.40	0.05	−0.40	0.05	−0.40	0.05	−0.42	0.05
REfrac	0.55	0.03	0.52	0.03	0.50	0.03	0.50	0.03	0.50	0.03
HMrate	−0.29	0.03	−0.29	0.04	−0.29	0.03	−0.29	0.04	−0.30	0.04
AltMinRate	2.08	0.04	2.04	0.04	2.02	0.04	2.04	0.04	2.07	0.04
EnergyRate	0.35	0.03	0.39	0.03	0.40	0.03	0.42	0.03	0.41	0.03
EITCrate	−0.60	0.02	−0.62	0.02	−0.64	0.02	−0.65	0.02	−0.67	0.02
UnempRate	0.28	0.02	0.32	0.02	0.34	0.01	0.34	0.02	0.34	0.02
lAvgAGI	0.22	0.00	0.24	0.00	0.24	0.00	0.25	0.00	0.25	0.00
lreturn	0.07	0.00	0.07	0.00	0.07	0.00	0.07	0.00	0.06	0.00
y11	0.02	0.00	0.01	0.00	0.01	0.00	0.00	0.00	−0.01	0.00
y12	0.03	0.00	0.03	0.00	0.02	0.00	0.02	0.00	0.01	0.00
y13	0.08	0.00	0.07	0.00	0.07	0.00	0.06	0.00	0.05	0.00
y14	0.10	0.00	0.10	0.00	0.09	0.00	0.08	0.00	0.07	0.00
y15	0.12	0.00	0.11	0.00	0.11	0.00	0.10	0.00	0.09	0.00
y16	0.12	0.00	0.12	0.00	0.12	0.00	0.12	0.00	0.11	0.00
σ	0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.00
γ	2.10	0.02	0.81	0.01	−0.01	0.01	−0.82	0.01	−2.08	0.02
φ^2	0.05	0.00	0.04	0.00	0.04	0.00	0.04	0.00	0.04	0.00

Table 10: Results from the REQ model for the entire US data–posterior mean (MEAN) and posterior standard deviation (SD) of the parameters.

	10th quantile		25th quantile		50th quantile		75th quantile		90th quantile	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
Intercept	5.69	0.02	5.67	0.02	5.65	0.02	5.75	0.03	5.85	0.03
SSBfrac	−0.75	0.02	−0.84	0.02	−0.95	0.02	−0.96	0.02	−0.89	0.02
FarmFrac	−0.43	0.05	−0.42	0.05	−0.40	0.05	−0.41	0.05	−0.44	0.05
REfrac	0.52	0.03	0.53	0.03	0.50	0.03	0.47	0.03	0.47	0.03
HMrate	−0.21	0.03	−0.25	0.03	−0.29	0.04	−0.25	0.04	−0.24	0.03
AltMinRate	2.11	0.04	2.04	0.04	2.02	0.04	2.02	0.04	2.00	0.04
EnergyRate	0.18	0.03	0.29	0.03	0.40	0.03	0.40	0.03	0.35	0.03
EITCrate	−0.56	0.02	−0.58	0.02	−0.64	0.02	−0.68	0.02	−0.69	0.02
UnempRate	0.20	0.02	0.26	0.02	0.33	0.02	0.34	0.02	0.36	0.02
lAvgAGI	0.18	0.00	0.21	0.00	0.24	0.00	0.24	0.00	0.24	0.00
lreturn	0.08	0.00	0.07	0.00	0.07	0.00	0.06	0.00	0.06	0.00
y11	0.03	0.00	0.02	0.00	0.01	0.00	−0.01	0.00	−0.02	0.00
y12	0.05	0.00	0.04	0.00	0.02	0.00	0.01	0.00	−0.01	0.00
y13	0.09	0.00	0.08	0.00	0.07	0.00	0.04	0.00	0.03	0.00
y14	0.12	0.00	0.11	0.00	0.09	0.00	0.07	0.00	0.05	0.00
y15	0.13	0.00	0.12	0.00	0.11	0.00	0.09	0.00	0.08	0.00
y16	0.13	0.00	0.13	0.00	0.12	0.00	0.11	0.00	0.10	0.00
σ	0.01	0.00	0.02	0.00	0.02	0.00	0.02	0.00	0.01	0.00
φ^2	0.05	0.00	0.05	0.00	0.04	0.00	0.05	0.00	0.05	0.00

References

Abrevaya, J. and Dahl, C. M. (2008), "The effects of birth inputs on birthweight: Evidence from quantile estimation on panel data," *Journal of Business & Economic Statistics*, 26, 379–397.

Alhamzawi, R. (2016), "Bayesian Model Selection in Ordinal Quantile Regression," *Computational Statistics and Data Analysis*, 103, 68–78.

Alhamzawi, R. and Ali, H. T. M. (2018), "Bayesian Quantile Regression for Ordinal Longitudinal Data," *Journal of Applied Statistics*, 45, 815–828.

Andrews, D. F. and Mallows, C. L. (1974), "Scale Mixtures of Normal Distributions," *Journal of the Royal Statistical Society – Series B*, 36, 99–102.

Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Kluwer Academic Publsihers, Dordrecht.

Antoniak, C. E. (1974), "Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems," *The Annals of Statistics*, 2, 1152–1174.

Beer, C. and Riedl, A. (2012), "Modelling Spatial Externalities in Panel Data: The Spatial Durbin Model Revisited," *Papers in Regional Science*, 91, 299–319.

Benhabib, J. and Bisin, A. (2018), "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, 56, 1261–1291.

Benoit, D. F. and Poel, D. V. D. (2010), "Binary Quantile Regression: A Bayesian Approach based on the Asymmetric Laplace Distribution," *Journal of Applied Econometrics*, 27, 1174–1188.

Botev, Z. I. (2016), "The Normal Law Under Linear Restrictions: Simulation and Estimation via Minimax Tilting," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 79, 125–148.

Bresson, G., Lacroix, G., and Rahman, M. A. (2021), "Bayesian Panel Quantile Regression for Binary Outcomes with Correlated Random Effects: An Application on Crime Recidivism in Canada," *Empirical Economics*, 60, 227–259.

Chamberlain, G. (1984), "Panel Data," in *Handbook of Econometrics*, eds. Z. Griliches and M. D. Intriligator, pp. 1247–1318, Amsterdam, North-Holland.

Chatterjee, S. and Eyigungor, B. (2015), "A Seniority Arrangement for Sovereign Debt," *American Economic Review*, 105, 3740–65.

Chen, X. and Tokdar, S. T. (2021), "Joint Quantile Regression for Spatial Data," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 83, 826–852.

Chib, S. (1995), "Marginal Likelihood from the Gibbs Output," *Journal of the American Statistical Association*, 90, 1313–1321.

Chib, S. and Carlin, B. P. (1999), "On MCMC Sampling in Hierarchical Longitudinal Models," *Journal of the Royal Statistical Society – Series B*, 9, 17–26.

Chib, S. and Greenberg, E. (1995), "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, 49, 327–335.

Chib, S. and Jeliazkov, I. (2001), "Marginal Likelihood from the Metropolis-Hastings Output," *Journal of the American Statistical Association*, 96, 270–281.

Chib, S. and Jeliazkov, I. (2006), "Inference in Semiparametric Dynamic Models for Binary Longitudinal Data," *Journal of the American Statistical Association*, 101, 685–700.

Chib, S. and Ramamurthy, S. (2010), "Tailored randomized block MCMC methods with application to DSGE models," *Journal of Econometrics*, 155, 19–38.

Das, P. and Ghosal, S. (2017), "Bayesian quantile regression using random B-spline series prior," *Computational Statistics & Data Analysis*, 109, 121–143.

Devroye, L. (2014), "Random Variate Generation for the Generalized Inverse Gaussian Distribution," *Statistics and Computing*, 24, 239–246.

Dunson, D. B. (2009), "Nonparametric Bayes local partition models for random effects," *Biometrika*, 96, 249–262.

Escobar, M. D. and West, M. (1995), "Bayesian Density Estimation and Inference Using Mixtures," *Journal of the American Statistical Association*, 90, 577–588.

Ferguson, T. S. (1973), "A Bayesian analysis of some nonparametric problems," *The Annals of Statistics*, 1, 209–230.

Galvao, A. F. (2011), "Quantile regression for dynamic panel data with fixed effects," *Journal of Econometrics*, 164, 142–157.

Galvao, A. F., Gu, J., and Volgushev, S. (2020), "On the unbiased asymptotic normality of quantile regression with fixed effects," *Journal of Econometrics*, 218, 178–215.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013), *Bayesian Data Analysis*, Chapman & Hall, New York.

Geraci, M. and Bottai, M. (2007), "Quantile Regression for Longitudinal Data Using the Asymmetric Laplace Distribution," *Biostatistics*, 8, 140–154.

Geraci, M. and Bottai, M. (2014), "Linear Quantile Mixed Models," *Statistics and Computing*, 24, 461–479.

Glaeser, E. L., Gyourko, J., and Saks, R. E. (2020), "Why have House Prices Gone Up?" *American Economic Review*, 95, 329–333.

Gneiting, T. and Raftery, A. E. (2007), "Strictly Proper Scoring Rules, Prediction, and Estimation," *Journal of the American Statistical Association*, 102, 359–378.

Goncalves, K. C. M., Migon, H. S., and Bastos, L. S. (2022), "Dynamic Quantile Linear Models: A Bayesian Approach," *Bayesian Analysis*, 15, 335–362.

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

Hanson, T. and Johnson, W. O. (2002), "Modeling Regression Error with a Mixture of Polya Trees," *Journal of the American Statistical Association*, 97, 1020–1033.

Hembre, E. and Dantas, R. (2022), "Tax Incentives and Housing Decisions: Effects of the Tax Cut and Jobs Act," *Regional Science and Urban Economics*, 95, 103800.

Ho, R. K. W. and Hu, I. (2008), "Flexible modelling of random effects in linear mixed models - A Bayesian approach," *Computational Statistics & Data Analysis*, 52, 1347–1361.

Jackson, K. (2018), "Regulation, Land Constraints, and California's Boom and Bust," *Regional Science and Urban Economics*, 68, 130–147.

Kato, K., F. Galvao, A., and Montes-Rojas, G. V. (2012), "Asymptotics for panel quantile regression models with individual effects," *Journal of Econometrics*, 170, 76–91.

Kobayashi, G. and Kozumi, H. (2012), "Bayesian Analysis of Quantile Regression for Censored Dynamic Panel Data Model," *Computational Statistics*, 27, 359–380.

Koenker, R. (2004), "Quantile regression for longitudinal data," *Journal of Multivariate Analysis*, 91, 74–89.

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

Koenker, R. and Bassett, G. (1978), "Regression Quantiles," *Econometrica*, 46, 33–50.

Koenker, R. and Machado, J. A. F. (1999), "Goodness of Fit and Related Inference Processes for Quantile Regression," *Journal of the American Statistical Association*, 94, 1296–1310.

Koop, G. and Tobias, J. L. (2004), "Learning about heterogeneity in returns to schooling," *Journal of Applied Econometrics*, 19, 827–849.

Kottas, A. and Gelfand, A. E. (2001), "Bayesian semiparametric median regression modeling," *Journal of the American Statistical Association*, 96, 1458–1468.

Kottas, A. and Krnjajić, M. (2009), "Bayesian Semiparametric Modelling in Quantile Regression," *Scandinavian Journal of Statistics*, 36, 297–319.

Kozumi, H. and Kobayashi, G. (2011), "Gibbs Sampling Methods for Bayesian Quantile Regression," *Journal of Statistical Computation and Simulation*, 81, 1565–1578.

Kyung, M., Gill, J., and Casella, G. (2010), "Estimation in Dirichlet random effects models," *The Annals of Statistics*, 38, 979–1009.

Lamarche, C. (2010), "Robust penalized quantile regression estimation for panel data," *Journal of Econometrics*, 157, 396–408.

Lamarche, C. (2021), "Quantile Regression for Panel Data and Factor Models," *Oxford Research Encyclopedia of Economics and Finance.*

Loewenstein, L. and Willen, P. (2023), "House Prices and Rents in the 21st Century," *NBER Working Paper 31013*.

Luo, Y., Lian, H., and Tian, M. (2012), "Bayesian Quantile Regression for Longitudinal Data Models," *Journal of Statistical Computation and Simulation*, 82, 1635–1649.

Maheshwari, P. and Rahman, M. A. (2023), "bqror: An R package for Bayesian Quantile Regression in Ordinal Models," *The R Journal*, 15, 39–55.

März, A., Klein, N., Kneib, T., and Musshoff, O. (2016), "Analysing Farmland Rental Rates using Bayesian Geoadditive Quantile Regression," *European Review of Agricultural Economics*, 43, 663–698.

Mundlak, Y. (1978), "On the Pooling of Time Series and Cross Section Data," *Econometrica*, 46, 69–85.

Nascimento, M. G. L. and Goncalves, K. C. M. (2021), "Bayesian Variable Selection in Quantile Regression with Random Effects: An Application to Municipal Human Development Index," *Journal of Applied Statistics*.

Neal, R. M. (2000), "Markov Chain Sampling Methods for Dirichlet Process Mixture Models," *Journal of Computational and Graphical Statistics*, 9, 249–265.

Ojha, M. and Rahman, M. A. (2021), "Do Online Courses Provide an Equal Educational Value Compared to In-Person Classroom Teaching? Evidence from U.S. Survey Data using Quantile Regression," *Education Policy Analysis Archives*, 29, 1–25.

Quigley, J. M. and Raphael, S. (2004), "Is Housing Unaffordable? Why Isn't It More Affordable?" *Journal of Economic Perspectives*, 18, 191–214.

Rahman, M. A. (2016), "Bayesian Quantile Regression for Ordinal Models," *Bayesian Analysis*, 11, 1–24.

Rahman, M. A. and Karnawat, S. (2019), "Flexible Bayesian Quantile Regression in Ordinal Models," *Advances in Econometrics*, 40B, 211–251.

Rahman, M. A. and Vossmeyer, A. (2019), "Estimation and Applications of Quantile Regression for Binary Longitudinal Data," *Advances in Econometrics*, 40B, 157–191.

Reich, B. J., Bondell, H. D., and Wang, H. J. (2010), "Flexible Bayesian quantile regression for independent and clustered data," *Biostatistics*, 11, 337–352.

Roberts, G. O. and Sahu, S. K. (1997), "Updating Schemes, Correlation Structure, Blocking and Parameterization for the Gibbs Sampler," *Journal of the Royal Statistical Society – Series B*, 59, 291—-317.

Ruppert, D., Wand, M. P., and Carroll, R. J. (2003), *Semiparametric Regression*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

Schwarz, G. E. (1978), "Estimating the Dimension of a Model," *Annals of Statistics*, 6, 461–464.

Shampine, L. F. (2008), “Vectorized adaptive quadrature in MATLAB,” *Journal of Computational and Applied Mathematics*, 211, 131–140.

Taddy, M. A. and Kottas, A. (2010), “A Bayesian Nonparametric Approach to Inference for Quantile Regression,” *Journal of Business & Economic Statistics*, 28, 357–369.

Thomschke, L. (2015), “Changes in the Distribution of Rental Prices in Berlin,” *Regional Science and Urban Economics*, 51, 88–100.

Tierney, L. (1994), “Markov Chains for Exploring Posterior Distributions,” *Annals of Statistics*, 22, 1701–1761.

Turek, D., de Valpine, P., Paciorek, C. J., and Anderson-Bergman, C. (2017), “Automated Parameter Blocking for Efficient Markov Chain Monte Carlo Sampling,” *Bayesian Analysis*, 12, 465–490.

Waltl, S. R. (2018), “Estimating Quantile-Specific Rental Yields for Residential Housing in Sydney,” *Regional Science and Urban Economics*, 68, 204–225.

Yan, Y. and Kottas, A. (2017), “A New Family of Error Distributions for Bayesian Quantile Regression,” *https://arxiv.org/abs/1701.05666.*

Yang, Y. and Tokdar, S. T. (2017), “Joint Estimation of Quantile Planes Over Arbitrary Predictor Spaces,” *Journal of the American Statistical Association*, 112, 1107–1120.

Yu, K. and Moyeed, R. A. (2001), “Bayesian Quantile Regression,” *Statistics and Probability Letters*, 54, 437–447.

Yu, Z., Yao, X., and Liu, H. (2025), “Bayesian Elastic Net Variable Selection and Application for Spatial Quantile Panel Autoregressive Model,” *Communications in Statistics–Simulation and Computation*, pp. 1–30.