Stigma and Spillovers in Emergency Lending: A Bayesian Quantile Regression Approach for Censored Longitudinal Data

Oncel Aldanmaz[1], Mohammad Arshad Rahman[2], and Angela Vossmeyer*[3]

[1]Meta
[2]Indian Institute of Technology Kanpur
[3]Claremont McKenna College & NBER

August 25, 2026

Abstract

Emergency lending programs often struggle to reach banks that fear participation may signal weakness. We study whether uptake of Reconstruction Finance Corporation assistance during the Great Depression depended on whether neighboring banks had recently participated. We estimate Bayesian quantile models for binary longitudinal data and develop a new censored quantile longitudinal model to capture both participation and intensity margins. Neighboring uptake positively predicts participation, especially where prior engagement was lowest. These patterns are consistent with a shift toward a pooling equilibrium, in which rising participation reduces the stigma of borrowing and increases familiarity with the program, together driving the spatial diffusion of emergency lending participation.

JEL classification: G21, C30, N12.

Keywords: Bayesian quantile regression; Censored data; Spatial spillovers; Bank stigma; Financial crises.

*The authors thank the Lowe Institute of Political Economy at Claremont McKenna College for support. The authors may be reached at oncelaldanmaz@meta.com, marshad@iitk.ac.in, and angela.vossmeyer@cmc.edu.

1 Introduction

During a financial crisis, governments often step in to lend directly to distressed banks, but the decision to receive assistance is complex. Emergency borrowing can reveal information about a bank's condition to depositors, counterparties, and regulators. If participation is read as a signal of weakness, many institutions may avoid assistance they could otherwise use, leaving the pool of visible participants skewed toward the most troubled firms and deepening the stigma attached to participation. Policymakers have long acknowledged this concern. Former Federal Reserve Chairman Ben Bernanke explained that banks avoided the discount window during the 2007–2008 crisis because reliance on it, if it became known, might lead market participants to infer weakness (Bernanke, 2009). The Federal Reserve responded by designing the Term Auction Facility, which explicitly masked individual participation within a large pool of auction winners. Empirical work has since shown that banks were willing to pay meaningfully higher rates simply to avoid the discount window's stigma and instead borrow from the Term Auction Facility (Armantier et al., 2015). Fiscal policymakers reasoned the same way: officials have recounted that the heads of nine of the largest U.S. financial firms were pressed to accept capital under the Troubled Asset Relief Program specifically to make the program appear less stigmatizing from the perspective of smaller institutions (Geithner, 2014). The logic behind these decisions is that stigma is not fixed and it depends on how many, and which, institutions are seen to participate. When participation is rare, it signals distress; when it becomes common, the signal weakens and the reputational cost of joining falls. This suggests that the decision to seek emergency assistance may be socially contagious: banks may become more willing to participate once they observe others, particularly their peers, doing so.

This paper tests that idea using the Reconstruction Finance Corporation (RFC), the bank rescue program during the Great Depression. We ask whether a bank's decision to seek RFC assistance was shaped by whether neighboring banks had recently done so. More specifically, we look at whether participation in an emergency lending program exhibits positive spatial spillovers. The object of interest is uptake itself: whether and how readily banks sought assistance, not the downstream economic consequences of receiving it. This question matters for three reasons. First, if participation decisions are contagious across space, then treating each institution's enrollment as an independent event may misrepresent how a program's reach actually expanded over time and

geography. Second, understanding that participation itself is contagious speaks directly to how emergency facilities should be designed and announced. Third, examining where in the distribution of participation propensity these spillovers are strongest reveals whether emergency lending programs can draw in reluctant or previously unengaged institutions through the observed behavior of their peers, or whether uptake instead remains concentrated among banks already predisposed to seek federal support

To investigate this question, we construct a novel county-level monthly panel spanning January 1932 through December 1935, built from digitized Reconstruction Finance Corporation card index records covering the universe of bank participation in RFC lending across the United States. The resulting dataset allows us to observe not just whether a bank sought assistance, but the presence and timing of participation for spatially adjacent banks. We estimate two complementary models. The first is the Bayesian quantile regression model for binary longitudinal data (QBLD), which characterizes how neighboring lending exposure shifts the probability of participation across the full conditional distribution of lending propensity. The second is a new Bayesian quantile regression model for censored longitudinal data (QCLD), which extends the QBLD framework to the continuous lending amount, accommodating the heavy left-censoring at zero that characterizes the RFC data. Together, the two models allow us to capture spillover effects both on the extensive margin (whether there is participation at all) and on the intensive margin (how much borrowing occurs).

Our approach is particularly appealing in this setting because RFC participation is both heavily skewed and heterogeneous across counties. Most county-months record no lending at all, while positive amounts range widely, so a single average spillover would misrepresent most counties in the panel. There is also no reason neighboring participation should matter equally across the distribution. A county already active in the program is in a different position than one that has never engaged with federal credit, and estimating a single mean effect ignores that difference. Quantile methods let us trace how the spillover varies across the full conditional distribution of lending rather than collapsing it to a single number, revealing not just whether neighboring participation matters but where it matters most. Capturing this kind of heterogeneity in a panel setting has posed well-known theoretical and computational challenges for frequentist quantile estimation, owing to the difficulty of removing latent unit-level heterogeneity; the Bayesian framework we develop addresses these challenges.

We find that neighboring lending exposure is positively associated with participation across a wide range of quantiles, specifications, and subsamples. The effect is strongest at the lower tail of the conditional distribution (i.e., regions least likely to receive assistance), precisely where stigma and informational barriers are highest and where peer participation would send a strong signal. We also find that the spillovers are the largest for institutions with the least prior familiarity with federal credit facilities.[MORE RESULTS].

The RFC is a useful setting for studying this question for several reasons. First, the RFC was stigmatized. Banks were initially reluctant to borrow for fear of signaling weakness, thus providing the condition under which we can study the potential erosion of reluctance. Second, the program was large and long-lived, extending assistance to institutions across the country over several years, which allows the pool of participants to evolve as more institutions engaged. Third, RFC lending decisions were largely insulated from political pressure to enroll particular institutions. Mason (2003) shows that allocations were not well explained by standard rent-seeking measures. Fourth, the banking system of the 1930s was highly localized, with most banks operating within a single county and serving local depositors and borrowers. This localization lets us define a bank's peers as its geographic neighbors directly, rather than inferring a network from behavior or affiliation, and to test for spillovers across adjacent counties cleanly.

This paper contributes to several strands of the literature. The first is the theoretical literature on stigma in emergency lending, which characterizes stigma as an equilibrium outcome that depends on which banks participate and how visible that participation is (Philippon and Skreta, 2012; Ennis and Weinberg, 2013; Ennis, 2018; Bajaj, 2018; Gorton and Ordonez, 2017). In particular, research has shown that participation could signal an inability to find funding elsewhere, resulting in a separating equilibrium in which participation and firm quality are linked. An implication is that greater uptake among institutions should weaken the signal, bringing in a pool of borrowers of varying quality. We empirically test this by looking at neighboring or peer institutions to predict how their uptake influences the probability of RFC participation among nearby banks.

The empirical literature on stigma includes Armantier et al. (2015), Anbil (2018), Vossmeyer (2019), Armantier and Holt (2020), Armantier and Holt (2025), Anbil and Vossmeyer (2019), Anbil and Vossmeyer (2021), Armantier et al. (2024), and Beyhaghi and Gerlach (2025). Armantier et al. (2015), Armantier et al. (2024), and Beyhaghi and Gerlach (2025) focus on the cost of stigma

by looking at the premium that banks are willing to pay to avoid stigma and instead approach an alternative, more expensive facility. Anbil (2018) and Vossmeyer (2019) study an unexpected stigma revelation where banks faced scrutiny from market participants for receiving emergency assistance. Anbil and Vossmeyer (2019) and Anbil and Vossmeyer (2021) subsequently show that when a facility is not stigmatized it attracts a larger pool of borrowers of varying quality, whereas stigmatized facilities attract weaker banks. We study a single stigmatized facility and ask whether its signal weakens as participation spreads across neighboring banks.

Another related literature has to do with barriers to entry into government programs. Anbil et al. (2023) show that banks already familiar with federal credit facilities were more likely to take up the Paycheck Protection Program Liquidity Facility, since prior experience lowered the information costs of engaging with a new one. Similar frictions appear well outside of banking. A substantial literature documents incomplete take-up of programs among eligible participants, attributing much of the shortfall to information, complexity, and administrative barriers, which diminish as participation increases (Bertrand et al., 2000; Currie, 2006; Dahl et al., 2014; Bhargava and Manoli, 2015; Finkelstein and Notowidigdo, 2019). We hypothesize that borrowing by neighboring banks worked in much the same way, lowering information barriers and increasing participation.

Our next contribution is methodological. We develop a censored panel quantile regression (QCLD) model that extends Bayesian quantile methods for panel data to censored outcomes, the structure that characterizes RFC lending amounts. Building on the asymmetric Laplace working likelihood (Yu and Moyeed, 2001) and its normal-exponential mixture representation (Kozumi and Kobayashi, 2011), we cast the model in a conditionally Gaussian hierarchical form that accommodates dynamic dependence and unit-specific effects. Extending the censored dynamic panel sampler of (Kobayashi and Kozumi, 2012), we develop an efficient Gibbs algorithm that jointly samples the regression coefficients and censored latent variables while integrating out the random effects, reducing the posterior correlation that slows the non-blocked sampler; the gains from such blocking are well documented (Chib and Carlin, 1999; Rahman and Vossmeyer, 2019). We further derive interpretable covariate effects and the marginal likelihood for model comparison (Chib, 1995a). This framework lets us move beyond the participation margin of the binary model to characterize how neighboring lending exposure shifts the full conditional distribution of lending amounts. The approach applies to any panel setting with a censored outcomes and quantile-varying effects.

2 Historical Background and Data

The Reconstruction Finance Corporation was established by an act of Congress signed on January 22, 1932, and began operations the following month. It was created to provide emergency assistance to a financial system, which was buckling under the weight of the Depression. Initially, the RFC extended collateralized loans to banks. Bank would apply to the RFC for assistance, the RFC would review the submitted application and assess the collateral quality, and then the RFC would determine whether to provide assistance and the amount. Loans carried interest rates set at or slightly above prevailing market rates (Mason, 2001).[1]

A July 1932 amendment to the RFC Act required that the names of borrowing institutions be reported to Congress, and these names were subsequently made public. Research has shown that this disclosure undermined the program because once participation became publicly observable, borrowing from the RFC could be read as a signal of distress. Banks grew reluctant to seek assistance for fear that appearing on the list would invite depositor runs (Anbil, 2018; Vossmeyer, 2019). This episode is central to our study, because participation is visible and stigmatizing, the behavior of nearby banks might shape a bank's own willingness to participate.

The RFC's mission broadened in 1933. Following the nationwide bank holiday in March, the Emergency Banking Act authorized the RFC to invest directly in the preferred stock of banks, shifting its support from collateralized loans toward recapitalization. Over the following years the RFC became a central instrument of New Deal financial policy, and it continued operating in various forms until the 1950s. Our analysis, however, focuses on the period from 1932 through 1935, when RFC assistance to banks was its main role and bank participation was a key concern to banks and policymakers alike. It is this decision, and its diffusion across space, that we study in this paper.

2.1 Data Sources and Construction

Our analysis combines four sources: RFC records, a directory of the universe of U.S. banks, census data, and a record of bank closures. The base of our panel is the 1932 edition of the *Rand McNally Bankers' Directory*, which lists every operating bank in the United States. This directory defines our panel of counties (counties in the panel must have contained at least one bank in the 1932

[1] A literature has studied the RFC's effectiveness, generally asking whether its assistance improved bank survival and financial conditions (Mason, 2001; Calomiris et al., 2013; Vossmeyer, 2016).

directory). It also classifies each bank as state- or nationally chartered, a distinction we exploit later in the paper. We also assign each bank to a Federal Reserve district using this data source.

The lending data come from the card index records of the Reconstruction Finance Corporation, held at the U.S. National Archives and first introduced to the literature by Vossmeyer (2016) and further described in Amujala et al. (2023). Each card documents the bank's name and location, the date of the RFC transaction, the type of assistance, the RFC's decision, and the dollar amount authorized. From the resulting records we construct, for each county and month, the dollar amount of RFC lending to banks. County population is taken from the 1930 U.S. Census. Because population is not observed at monthly frequency, we treat it as a time-invariant county characteristic over the sample period, assigning each county its 1930 value.

Next, we compile a record of discontinued banks from the 1934 and 1939 editions of the *Rand McNally Bankers' Directory*. These issues contain a "Discontinued Bank List," cataloging banks that had ceased operation, along with the manner and approximate timing of their closure. Matching these lists back to the 1932 bank directory allows us to identify which banks closed and when, from which we construct monthly county-level closure measures.

We aggregate the four sources into a single county-level monthly panel spanning 1932 to 1935. We assign each bank to a county using U.S. Census FIPS codes, and for each county and month, we then construct measures of total RFC lending to banks and total bank closures. To capture the spatial structure central to our question, we construct a county adjacency matrix from U.S. Census TIGER shapefiles using Queen contiguity, under which two counties are neighbors if they share either a border or a vertex. Applying this matrix to the own-county measures, we compute, for each county and month, the corresponding totals across all neighboring counties: neighboring RFC lending and neighboring bank closures. The adjacency structure is binary and unweighted, so each neighbor contributes equally. In our setting these geographic neighbors serve as economic peers. Banking in the 1930s was highly localized, with most banks operating within a single county and serving local depositors and borrowers, so nearby counties are the natural unit through which information about the RFC program and the behavior of participating banks would have traveled. Our interest is in whether this neighboring activity influenced a county's own uptake of RFC assistance.

The neighboring closure measures play an important role. A positive association between neigh-

boring and own-county participation could reflect genuine spillovers, but it could also arise if adjacent counties were simply exposed to the same regional shocks. Bank closures provide a direct measure of local distress. By conditioning on both own-county and neighboring closures, we absorb much of the common variation in financial conditions across nearby counties, helping to separate spillovers in participation from the shared shocks that would otherwise confound them.

The panel structure allows us to examine how spillovers vary across economically meaningful subsamples, defined by population size, bank charter type, and Federal Reserve district. Variation along each dimension is informative for a distinct reason. Charter type matters because national banks had direct access to the Federal Reserve's discount window, whereas state banks generally relied on the RFC as their primary source of emergency support, so the two may have differed in how strongly neighboring participation shaped their own. Federal Reserve districts may matter because monetary policy varied across regions. And population size matters because it shapes both the demand for credit and the visibility of any single bank's participation. We exploit each of these dimensions in econometric analyses.

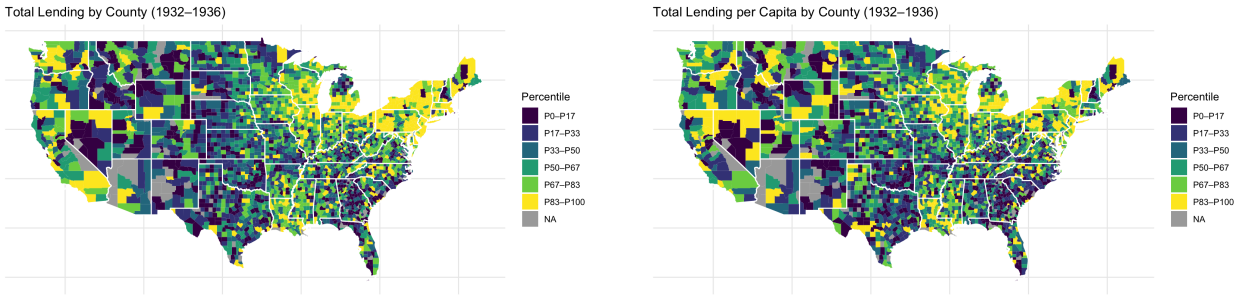
2.2 Descriptive Statistics

The final panel comprises 2,979 counties observed monthly from 1932 to 1935. Two features of the data shape the modeling choices that follow and preview the spillover patterns we later estimate.

The first is the participation rate. Only 12.7% of county-month observations record any positive RFC lending to banks; the remaining 87.3% are zeros. This heavy concentration at zero is the defining feature of the outcome and motivates our two-pronged empirical approach. Whether a county participated at all is itself the object of interest, which the binary model of Section 3 is designed to capture, while the censored model of Section 4 additionally accommodates the large mass at zero when modeling the amount of lending. Notably, neighboring exposure is far more common than own-county participation with 43.9% of county-month observations having at least one adjacent county that received lending in the same month. Variation in the behavior of a county's neighbors is therefore widespread.

The second is that participation, while dispersed across the U.S., had some clustering. Figure 1 maps cumulative RFC lending by county over 1932–1935, in both total and per-capita terms. Lending is visibly concentrated in the Northeast and Midwest. While the per capita normalization

redistributes some of the activity, the clustering remains.



(a) Total lending by county

(b) Total lending per capita by county

Figure 1: Geographic distribution of RFC total lending, 1932–1935. Counties are binned by percentile of cumulative lending.

Table 1 describes our data and Figure 2 shows the share of RFC participating counties in each month-year and cumulative over time. While the monthly participation rate fluctuates with program expansions—rising sharply after the RFC opened in 1932 and again around the 1933–1934 recapitalization—the cumulative share of counties that had ever participated climbs to roughly 87% by the end of the sample. The contrast between the rarity of participation at any single point and its prevalence over time and across space is what makes the spillover question meaningful and estimable.

Table 1: Summary Statistics

Variable	Mean	Std. Dev.
County participation (any bank)	0.127	0.333
Adjacent participation	0.439	0.496
RFC lending to banks, own county	25.9	624.5
RFC lending to banks, adjacent counties	148.3	1,430.6
Own-county closures	0.034	0.249
Adjacent-county closures	0.205	0.704
Population (1930 Census)	40,185	136,241
Banks in county (1932)	6.163	7.462

Notes: The panel covers 2,979 counties observed monthly from January 1932 through December 1935 (142,992 county-month observations). Lending amounts are in thousands of dollars.

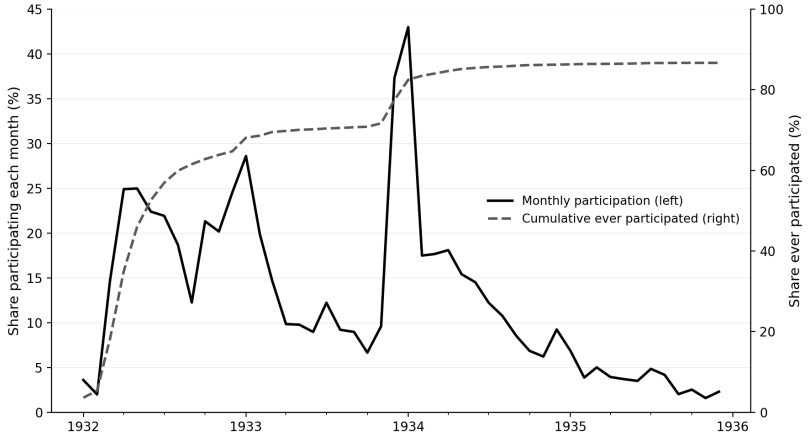


Figure 2: RFC participation over time, 1932–1935. The solid line (left axis) plots the share of counties whose banks received positive RFC lending in each month. The dashed line (right axis) plots the cumulative share of counties that had participated at least once by that month.

These data are employed in two econometric specifications. In the binary model (Section 3), the outcome is an indicator for whether a county received any RFC assistance in a given month. In the censored model (Section 4), the outcome is the amount of assistance received, left-censored at zero. These outcomes are regressed on a lagged measure of neighboring county participation together with the county's own lagged participation, the number of own-county and adjacent-county bank closures, county-specific random effects, and a full set of time fixed effects.

We enter all regressors at a one-month lag to mitigate simultaneity where a county's participation and its neighbors' participation are plausibly determined together within a month. Thus, lagging frames the estimand in terms of whether prior neighboring activity predicts subsequent own-county uptake, rather than capturing contemporaneous co-movement. The closure counts control for local financial distress, which helps separate regional shocks and common distress from spillovers. The county-specific random effect absorbs time-invariant heterogeneity, which is particularly valuable in this historical setting. County-month covariates are sparse in the 1930s, yet many of the local features that plausibly shaped RFC participation– banking regulation, industrial composition, population, and the structure of the local banking network – changed little in our higher-frequency time series. Finally, the time fixed effects, one per year-month, absorb aggregate shocks common to all counties, including macroeconomic conditions, seasonal patterns, and program-wide changes in RFC policy, so that identification comes from variation across counties within a given month rather than from the national trajectory of the program.

3 Quantile Regression for Binary Longitudinal Data (QBLD)

We first model the extensive margin of program participation: whether a county's banks received any RFC assistance in a given month. For this binary outcome we adopt the Bayesian quantile regression model for binary longitudinal data (QBLD) of Rahman and Vossmeyer (2019), which pairs an asymmetric Laplace working likelihood with a panel random-effects structure and delivers quantile-specific inference on a binary response. Because the model and its estimation are established, we describe them only briefly and refer the reader to Rahman and Vossmeyer (2019) for a full treatment.

Let $y_{it} \in \{0, 1\}$ indicate whether county i received any positive RFC lending to banks in month t, and let $p \in (0, 1)$ denote the quantile of interest. The QBLD model represents this outcome through a latent variable z_{it},

$$\begin{aligned} z_{it} &= x_{it}'\beta_p + s_{it}'\alpha_{ip} + \varepsilon_{it}, \qquad \forall\, i = 1, \ldots, n, \;\; t = 1, \ldots, T, \\ y_{it} &= \mathbb{1}\{z_{it} > 0\}, \end{aligned} \tag{1}$$

where x_{it} is a vector of covariates with common coefficients β_p, s_{it} is a vector of covariates with unit-specific effects α_{ip}, and ε_{it} follows an asymmetric Laplace distribution. In our application, x_{it} contains the lagged indicator of neighboring participation that is our coefficient of interest, along with the county's own lagged participation, lagged own-county and adjacent-county bank closures, and time fixed effects, while s_{it} contains a county-level intercept, so that α_{ip} is a scalar random effect absorbing time-invariant county heterogeneity. We assume $\alpha_{ip} \mid \Psi_p \sim N(0, \Psi_p)$ and suppress the quantile subscript p hereafter, with the understanding that all parameters remain quantile-specific.

The binary outcome is related to the latent variable through the sign restriction in (1): $y_{it} = 1$ when $z_{it} > 0$ and $y_{it} = 0$ otherwise. The asymmetric Laplace error admits the normal-exponential mixture representation (Kozumi and Kobayashi, 2011), which renders the model conditionally Gaussian and yields tractable full conditional distributions.

Estimation follows the blocked Gibbs sampler of Rahman and Vossmeyer (2019), in which the latent utilities z_{it} are drawn subject to the sign restriction implied by y_{it} and the remaining parameters are updated from their full conditionals; we implement it through the `qbld` package in R. We place a multivariate normal prior on β and standard conjugate priors on the remaining parameter.

4 Quantile Regression for Censored Longitudinal Data (QCLD)

4.1 Model

This section presents the censored quantile regression model for longitudinal data (QCLD). The model accommodates dynamic dependence by including lagged response values among the covariates (Kobayashi and Kozumi, 2012). We then develop an efficient Gibbs sampling algorithm based on parameter blocking, derive the posterior quantities of interest, including covariate effects, and describe the computation of the marginal likelihood for Bayesian model comparison.

Let y_{it} denote the response variable for unit i at time period t, left-censored at zero, and let $p \in (0,1)$ denote the quantile of interest. The QCLD model can be expressed conveniently through the latent variable z_{it} as

$$\begin{aligned} z_{it} &= x_{it}'\beta_p + s_{it}'\alpha_{ip} + \varepsilon_{it}, \qquad \forall\, i = 1,\ldots,n,\;\; t = 1,\ldots,T_i, \\ y_{it} &= \begin{cases} z_{it} & \text{if } y_{it} > 0, \\ 0 & \text{if } z_{it} \le 0, \end{cases} \end{aligned} \tag{2}$$

where $z_{it} = y_{it}$ when the response is positive (i.e., $y_{it} > 0$) and $z_{it} \le 0$ when it is left-censored at zero.[2]

Here x_{it} is a $k \times 1$ vector of explanatory variables and β_p is the corresponding $k \times 1$ vector of common regression coefficients, while s_{it} is an $l \times 1$ vector of covariates with unit-specific effects and α_{ip} is the associated $l \times 1$ vector of unit-specific parameters, with $E(\alpha_{ip}) = 0$ imposed for identification. To construct a working likelihood (Yu and Moyeed, 2001), the error term ε_{it} is assumed independently and identically distributed (*iid*) as an asymmetric Laplace (AL) distribution, denoted $AL(0, \sigma_p, p)$, where σ_p is the scale parameter (Yu and Zhang, 2005). Under this specification, the p-th conditional quantile of ε_{it} is zero, that is, $Q_{\varepsilon_{it}}(p \mid x_{it}, \alpha_{ip}) = 0$. Consequently, $z_{it} \mid \alpha_{ip} \sim AL(x_{it}'\beta_p + s_{it}'\alpha_{ip},\, \sigma_p,\, p)$ for all i and t, so that $Q_{z_{it}}(p \mid x_{it}, \alpha_{ip}) = x_{it}'\beta_p + s_{it}'\alpha_{ip}$.

We assume that the unit-specific effects satisfy $\alpha_{ip} \mid \Psi_p \sim N(0_l, \Psi_p)$ for $i = 1,\ldots,n$, where N denotes the normal distribution, though alternative specifications have also been considered in the literature. The QCLD model in equation (2) is general enough to accommodate both dynamic dependence and correlated random effects (Mundlak, 1978; Chamberlain, 1984). Dynamic dependence is introduced by including lagged values of the response variable in x_{it}, as in our application,

[2]Formally, define two index sets, $O_i = \{t : y_{it} > 0\}$ and $C_i = \{t : y_{it} = 0\}$. For $t \in O_i$, the latent variable satisfies $z_{it} = y_{it}$ and is therefore fully determined by the observed response. For $t \in C_i$, the latent variable is unobserved and must be inferred subject to the constraint $z_{it} \le 0$.

yielding a censored dynamic panel quantile regression model. Correlated random effects are accommodated by specifying the unit-specific effects as functions of a subset of the explanatory variables (Kobayashi and Kozumi, 2012; Bresson et al., 2021). Henceforth, for notational convenience, we suppress the quantile subscript p throughout the remainder of the paper, while emphasizing that all model components—including the regression coefficients, unit-specific effects, and covariance matrices—remain quantile-specific.

Because the AL density does not yield tractable conditional distributions, we adopt its normal-exponential mixture representation (Kozumi and Kobayashi, 2011) and express the error as

$$\begin{aligned} \varepsilon_{it} &= \sigma\theta w_{it} + \sigma\tau\sqrt{w_{it}}\, u_{it}, \qquad\qquad \forall\ i = 1, \ldots, n;\ t = 1, \ldots, T_i, \\ &= \theta\nu_{it} + \tau\sqrt{\sigma\nu_{it}}\, u_{it}, \end{aligned} \tag{3}$$

where $u_{it} \sim N(0,1)$ is mutually independent of $w_{it} \sim \mathcal{E}(1)$, and $\nu_{it} = \sigma w_{it} \sim \mathcal{E}(\sigma)$ so that $E(\nu_{it}) = \sigma$, with $\mathcal{E}$ denoting the exponential distribution. The constants are $\theta = \frac{1-2p}{p(1-p)}$ and $\tau = \sqrt{\frac{2}{p(1-p)}}$. This mixture representation has become the standard approach for Bayesian quantile regression and has been used widely to construct Markov chain Monte Carlo (MCMC) algorithms for a range of quantile models, including ordinal (Rahman, 2016), endogenous Tobit (Kobayashi, 2017), and dynamic linear specifications (Goncalves et al., 2022).

Substituting the normal-exponential mixture representation (3) into the QCLD model (2) yields

$$\begin{aligned} z_{it} &= x_{it}'\beta + s_{it}'\alpha_i + \theta\nu_{it} + \tau\sqrt{\sigma\nu_{it}}\, u_{it}, \qquad\qquad \forall\ i = 1, \ldots, n,\ \ t = 1, \ldots, T_i, \\ y_{it} &= \begin{cases} z_{it} & \text{if } y_{it} > 0, \\ 0 & \text{if } z_{it} \leq 0. \end{cases} \end{aligned} \tag{4}$$

It follows from equation (4) that $z_{it} \mid \alpha_i, \nu_{it} \sim N(x_{it}'\beta + s_{it}'\alpha_i + \theta\nu_{it},\, \tau^2\sigma\nu_{it})$ for all i and t. The hierarchical representation therefore transforms the asymmetric Laplace model into a conditionally Gaussian one. In doing so, the model inherits the tractability of the normal distribution, which allows us to construct an efficient Gibbs sampling algorithm for Bayesian estimation and posterior inference.

To complete the Bayesian specification, we assign prior distributions to the regression coefficients, the scale parameter, and the covariance matrix of the unit-specific effects,

$$\beta \sim N_k(\beta_0, B_0), \qquad \sigma \sim IG(n_0/2, d_0/2), \qquad \Psi \sim IW(r_0, R_0), \tag{5}$$

where IG and IW denote the inverse-Gamma[3] and inverse-Wishart[4] distributions, respectively. The normal prior for β provides a flexible specification for the regression coefficients and implies an ℓ_2 penalty on the quantile loss function (Yuan and Yin, 2010; Luo et al., 2012), while the IG prior ensures positivity of the scale parameter σ. The IW prior guarantees that the covariance matrix Ψ is positive definite and is the standard conjugate prior for covariance matrices in multivariate normal models. Together, these priors facilitate posterior computation by yielding tractable full conditional distributions within the Gibbs sampler.

4.2 Estimation

By Bayes' theorem, the "complete-data posterior" density is proportional to the product of the likelihood implied by the QCLD model (4) and the prior distributions (5),

$$\begin{aligned}
\pi(z,\beta,\alpha,\nu,\sigma,\Psi|y) \propto & \left\{\prod_{i=1}^{n}\prod_{t=1}^{T_i}\Big[I(y_{it}=z_{it})I(y_{it}>0)+I(y_{it}=0)I(z_{it}\le 0)\Big]\right\} \\
& \times\left\{\prod_{i=1}^{n}|\Lambda_i|^{-\frac{1}{2}}\right\}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left[(z_i-X_i\beta-S_i\alpha_i-\theta\nu_i)'\,\Lambda_i^{-1}\,(z_i-X_i\beta-S_i\alpha_i-\theta\nu_i)\right]\right\} \\
& \times\sigma^{-(\sum_{i=1}^{n}T_i)}\,\exp\left\{-\frac{1}{\sigma}\sum_{i=1}^{n}\sum_{t=1}^{T_i}\nu_{it}\right\}\times\exp\left\{-\frac{1}{2}(\beta-\beta_0)'B_0^{-1}(\beta-\beta_0)\right\}\times|\Psi|^{-\frac{n}{2}} \\
& \times\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\alpha_i'\Psi^{-1}\alpha_i\right\}\times\sigma^{-(\frac{n_0}{2}+1)}\exp\left\{-\frac{d_0}{2\sigma}\right\}\times|\Psi|^{-\left(\frac{r_0+l+1}{2}\right)}\exp\left\{-\frac{1}{2}\mathrm{tr}(\Psi^{-1}R_0)\right\},
\end{aligned} \tag{6}$$

where, for computational convenience, we stack the variables for each unit i over its time periods as $z_i=(z_{i1},\cdots,z_{iT_i})'$, $X_i=(x_{i1}',\cdots,x_{iT_i}')'$, $S_i=(s_{i1}',\cdots,s_{iT_i}')'$, $\nu_i=(\nu_{i1},\cdots,\nu_{iT_i})'$, $\Lambda_i^{1/2}=\mathrm{diag}(\tau\sqrt{\sigma\nu_{i1}},\cdots,\tau\sqrt{\sigma\nu_{iT_i}})$, and $u_i=(u_{i1},\cdots,u_{iT_i})'$.

The joint posterior in (6) has no closed-form expression, which necessitates MCMC methods for estimation and inference. In designing an MCMC algorithm for the QCLD model, a trade-off arises between *computational simplicity* and *sampling efficiency*. The first approach, which we refer to as the *non-blocked algorithm*, is computationally simpler but less efficient at sampling. The second,

[3]Suppose $\sigma\sim IG(n_0/2,d_0/2)$. The corresponding inverse-Gamma probability density function (*pdf*) is $f(\sigma)\propto\sigma^{-(n_0/2+1)}\exp\big(-d_0/2\sigma\big)$, defined on the support $\sigma>0$. The mean and variance are $E(\sigma)=d_0/(n_0-2)$ and $\mathrm{Var}(\sigma)=2d_0^2/\big[(n_0-2)^2(n_0-4)\big]$, which exist for $n_0>2$ and $n_0>4$, respectively. This corresponds to the shape-scale parameterization with shape $\alpha=n_0/2$ and scale $\beta=d_0/2$ (Gelman et al., 2013).

[4]Suppose $\Psi\sim IW(r_0,R_0)$. The inverse-Wishart *pdf* is $f(\Psi)\propto|\Psi|^{-(r_0+l+1)/2)}\exp\big(-\frac{1}{2}\mathrm{tr}(\Psi^{-1}R_0)\big)$, where $r_0\ge l$ and both Ψ and R_0 are $l\times l$ positive definite matrices. The mean $E(\Psi)=R_0/(r_0-l-1)$ exists for $r_0>l+1$. See Greenberg (2012) or Gelman et al. (2013) for further details.

which we propose, is computationally more demanding but achieves substantially greater sampling efficiency.

The non-blocked algorithm, presented as Algorithm 2 in Appendix A, modifies the Gibbs sampler of Kobayashi and Kozumi (2012). Although straightforward to derive and implement, it can mix poorly because of the strong posterior dependence between (β, α) and (z, α). This dependence is especially pronounced when $s_{it} \subseteq x_{it}$, that is, when the covariates associated with the unit-specific effects form a subset of the regression covariates, as is typically the case. Updating these quantities conditionally on one another then produces highly correlated Gibbs draws, resulting in slow mixing of the Markov chain (Chib and Carlin, 1999).

The proposed *blocked algorithm*,[5] presented in Algorithm 1, exploits a blocking technique in which highly correlated parameters and variables are updated jointly to improve the mixing of the Markov chain. This strategy reduces the correlation in the MCMC draws and increases sampling efficiency, at the expense of greater computational complexity, particularly in sampling the latent variable z. The benefits for MCMC efficiency nonetheless outweigh this additional burden. The advantages of blocked sampling are well documented by Chib and Carlin (1999) and Chib and Jeliazkov (2006), and more recently in the panel quantile regression literature by Rahman and Vossmeyer (2019) and Bresson et al. (2021).

The blocked update begins by integrating out the random effects α and jointly sampling the regression coefficients β and latent variables z. Specifically, β is first drawn from a multivariate normal distribution, after which the censored components of each z_i are drawn from a truncated multivariate normal distribution using the Gibbs sampler of Geweke (1991), which sequentially updates each component from its conditional univariate truncated normal distribution; see Appendix B for details. Conditional on these draws, the random effects α_i are sampled from an updated multivariate normal distribution. The remaining quantities are then updated with standard Gibbs steps: the latent weights ν are sampled element-wise from generalized inverse Gaussian (GIG) distributions, the scale parameter σ from an updated inverse-Gamma distribution, and the covariance matrix Ψ from an updated inverse-Wishart distribution.

We conclude this section by noting that the proposed QCLD model and its associated algorithms are developed for the general setting in which heterogeneity is present in both the intercept

[5]The derivations of the conditional posterior distributions are provided in Appendix B.

Algorithm 1 (Blocked Sampling)

---

1. Sample (β, z), marginally of α, in one block as in the following two sub-steps.

   (a) Let $\Lambda_i^{1/2} = D_{\tau\sqrt{\sigma\nu_i}}$ and $\Omega_i = (S_i \Psi S_i' + \Lambda_i)$. Sample $[\beta | z, \nu, \sigma, \Psi] \sim N(\tilde{\beta}, \tilde{B})$, where,

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

   (b) Sample $[z_i | y_i, \beta, \nu_i, \sigma, \Psi] \sim TMVN_{B_i}(X_i\beta + \theta\nu_i, \Omega_i)$ for all $i = 1, \ldots, n$, where, $TMVN$ denotes a truncated multivariate normal distribution and $B_i = (B_{i1} \times B_{i2} \times \cdots \times B_{iT_i})$ denotes the truncation region. For $t = 1, \ldots, T_i$, the set B_{it} is the singleton $\{y_{it}\}$ when $y_{it} > 0$, and equals the interval $(-\infty, 0]$ when $y_{it} = 0$. Define two index sets: $O_i = \{t : y_{it} > 0\}$ and $C_i = \{t : y_{it} = 0\}$ and partition the latent vector, mean vector, and covariance matrix into observed and censored components:

$$z_i = \begin{pmatrix} z_i^c \\ z_i^o \end{pmatrix}, \qquad \mu_i = X_i\beta + \theta\nu_i = \begin{pmatrix} \mu_i^c \\ \mu_i^o \end{pmatrix}, \qquad \Omega_i = \begin{pmatrix} \Omega_i^{cc} & \Omega_i^{co} \\ \Omega_i^{oc} & \Omega_i^{oo} \end{pmatrix},$$

   where the superscripts o and c denote the observed and censored components, respectively. Since $z_i^o = y_i^o$ is observed exactly, only z_i^c requires sampling. Its conditional distribution is,

$$\begin{aligned} [z_i^c | z_i^o, \beta, \nu_i, \sigma, \Psi] &\sim TMVN_{(-\infty,0]^{|C_i|}}\Big(\eta_i^c,\, \Sigma_i^c\Big), \qquad \text{where} \\ \eta_i^c &= \mu_i^c + \Omega_i^{co}(\Omega_i^{oo})^{-1}(z_i^o - \mu_i^o), \qquad \text{and} \\ \Sigma_i^c &= \Omega_i^{cc} - \Omega_i^{co}(\Omega_i^{oo})^{-1}\Omega_i^{oc}, \end{aligned}$$

   are the conditional mean and conditional variance, respectively. In addition, $TMVN_{(-\infty,0]^{|C_i|}}(\cdot,\cdot)$ denotes a multivariate normal distribution truncated to $(-\infty, 0]^{|C_i|}$, where $|C_i|$ is the number of censored variables for the i-th unit. Samples of z_i^c are obtained using the Gibbs sampler of Geweke (1991), which sequentially updates each component from its conditional univariate truncated normal distribution. This is explained in Appendix B.

2. Sample $[\alpha_i | z, \beta, \nu, \sigma, \Psi] \sim N(\tilde{a}, \tilde{A})$ for $i = 1, \ldots, n$, where,

$$\tilde{A}^{-1} = \left(S_i' \, \Lambda_i^{-1} \, S_i + \Psi^{-1} \right) \quad \text{and} \quad \tilde{a} = \tilde{A} \left(S_i' \Lambda_i^{-1} \left(z_i - X_i\beta - \theta\nu_i \right) \right).$$

3. Sample $[\nu_{it} | z_{it}, \beta, \alpha_i, \sigma] \sim GIG(0.5, \tilde{\lambda}_{it}, \tilde{\eta})$ for $i = 1, \ldots, n$, and $t = 1, \ldots, T_i$, where,

$$\tilde{\lambda}_{it} = \left(\frac{z_{it} - x_{it}'\beta - s_{it}'\alpha_i}{\tau\sqrt{\sigma}} \right)^2 \quad \text{and} \quad \tilde{\eta} = \left(\frac{\theta^2}{\tau^2\sigma} + \frac{2}{\sigma} \right).$$

4. Sample $[\sigma | z, \beta, \alpha, \nu] \sim IG(\tilde{n}/2, \tilde{d}/2)$, where

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

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

---

and a subset of the slope coefficients, often referred to as multivariate heterogeneity. In many empirical applications, however, including ours, heterogeneity enters only through the intercept. This special case is accommodated within the proposed algorithms without any conceptual modification. When Ψ is scalar, the inverse-Wishart distribution reduces to an inverse-Gamma distribution, so we may continue to use the prior for Ψ in equation (5) together with the full conditional distributions in Step 5 of Algorithms 1 and 2. Alternatively, we may specify the equivalent prior $\Psi \sim IG\left(n_1/2, d_1/2\right)$ and replace the corresponding full conditionals in Algorithms 1 and 2 with $\Psi \mid y, \alpha \sim IG\left(\tilde{n}_1/2, \tilde{d}_1/2\right)$, where $\tilde{n}_1 = n + n_1$ and $\tilde{d}_1 = \sum_{i=1}^{n} \alpha_i^2 + d_1$. No other modifications are required.

4.3 Posterior Quantities of Interest

Having estimated the QCLD model, the quantities of interest are the expected value of the response variable at a given quantile, evaluated either conditionally on the response being uncensored or unconditionally over the entire sample (i.e., both censored and uncensored observations). A further objective is to estimate the corresponding covariate (marginal) effects, which quantify how changes in the explanatory variables affect each of these expected values.

Exploiting the conditionally Gaussian representation of the QCLD model, the expectation of the response among uncensored observations, conditional on (α_i, ν_{it}), is

$$E(y_{it}|y_{it} > 0, \alpha_i, \nu_{it}) = \eta_{it} = \mu_{it} + \tau\sqrt{\sigma\nu_{it}}\,\frac{\phi(g_{it})}{\Phi(g_{it})}, \tag{7}$$

where $\mu_{it} = x_{it}'\beta + s_{it}'\alpha_i + \theta\nu_{it}$, $g_{it} = \frac{\mu_{it}}{\tau\sqrt{\sigma\nu_{it}}}$, and $\phi(\cdot)$ and $\Phi(\cdot)$ denote the probability density function (*pdf*) and cumulative distribution function (*cdf*) of the standard normal distribution, respectively. The expectation of the response over the entire sample is then

$$E(y_{it}|\alpha_i, \nu_{it}) = \Pr(y_{it} > 0)E(y_{it}|y_{it} > 0, \alpha_i, \nu_{it}) = \Phi(g_{it})\;\eta_{it}. \tag{8}$$

To obtain the marginal expectations $E(y_{it}|y_{it} > 0)$ and $E(y_{it})$, the latent mixture variable ν_{it} and the unit-specific effects α_i are integrated out with respect to their posterior distribution. These integrals are analytically intractable but are readily approximated by averaging over the posterior MCMC draws,

$$\widehat{E}(y_{it} \mid y_{it} > 0) = \frac{1}{M}\sum_{m=1}^{M} \eta_{it}^{(m)} \quad \text{and} \quad \widehat{E}(y_{it}) = \frac{1}{M}\sum_{m=1}^{M} \Phi\left(g_{it}^{(m)}\right)\,\eta_{it}^{(m)}, \tag{9}$$

where M is the number of post-burn-in MCMC draws, and $\eta_{it}^{(m)}$, $\mu_{it}^{(m)}$, and $g_{it}^{(m)}$ are obtained by substituting the m-th posterior draw of $(\beta, \alpha_i, \nu_{it}, \sigma)$ into their respective definitions.

We compute the average covariate effects analogously, by first evaluating the effect of a covariate on the expected value of y_{it}, conditional on (α_i, ν_{it}), and then averaging over the posterior MCMC draws and the sample. Suppose the j-th continuous covariate appears in x_{it} with coefficient β_j and in s_{it} with unit-specific coefficient α_{il}. The covariate effects on the conditional and unconditional expectations are

$$\frac{\partial E(y_{it}|y_{it}>0,\alpha_i,\nu_{it})}{\partial x_{it,j}} = (\beta_j+\alpha_{il})\left[1-\frac{\phi(g_{it})}{\Phi(g_{it})}\left(g_{it}+\frac{\phi(g_{it})}{\Phi(g_{it})}\right)\right], \tag{10}$$

$$\frac{\partial E(y_{it}|\alpha_i,\nu_{it})}{\partial x_{it,j}} = (\beta_j+\alpha_{il})\Phi(g_{it}). \tag{11}$$

The average covariate effects are obtained by averaging these expressions over the posterior draws of $(\beta, \alpha, \nu, \sigma)$ and over either the uncensored sample or the entire sample, depending on the quantity of interest. If the covariate appears only in x_{it}, the corresponding expressions follow by replacing $(\beta_j+\alpha_{il})$ with β_j, with everything else unchanged. Our computation of the average covariate effects accounts for both sampling and estimation uncertainty; the benefits of accounting for the latter, particularly in nonlinear models, are well demonstrated by Jeliazkov and Vossmeyer (2018).

For a finite change in a continuous variable, or for a discrete variable (e.g., binary or count), we instead compute the expected outcome under two covariate settings, holding all other covariates fixed, and take their difference. Specifically, suppose the j-th covariate takes the values a and b, giving rise to the covariate vectors $x_{it}^q = (x_{it,j}^q, x_{it,-j})$ and $s_{it}^q = (s_{it,l}^q, s_{it,-l})$ for $q = b, a$. The changes in the conditional (C) and unconditional (U) expectations are

$$\Delta_{y_{it}}^C = \eta_{it}^b - \eta_{it}^a = \left[\mu_{it}^b - \mu_{it}^a\right] + \tau\sqrt{\sigma\nu_{it}}\left[\frac{\phi(g_{it}^b)}{\Phi(g_{it}^b)} - \frac{\phi(g_{it}^a)}{\Phi(g_{it}^a)}\right], \tag{12}$$

$$\Delta_{y_{it}}^U = \Phi(g_{it}^b)\eta_{it}^b - \Phi(g_{it}^a)\eta_{it}^a = \left[\Phi(g_{it}^b)\mu_{it}^b - \Phi(g_{it}^a)\mu_{it}^a\right] + \tau\sqrt{\sigma\nu_{it}}\left[\phi(g_{it}^b) - \phi(g_{it}^a)\right], \tag{13}$$

where $\mu_{it}^q = (x_{it,j}^q, x_{it,-j})'(\beta_j, \beta_{-j}) + (s_{it,l}^q, s_{it,-l})'(\alpha_{il}, \alpha_{i,-l}) + \theta\nu_{it}$ and $g_{it}^q = \mu_{it}^q/(\tau\sqrt{\sigma\nu_{it}})$ for $q = b, a$. As before, the average covariate effects are obtained by averaging the expressions in (12) and (13) over the posterior MCMC draws and over the uncensored sample or the full sample, respectively.

4.4 Model Comparison

In the Bayesian framework, competing models are compared using the marginal likelihood, obtained by integrating the likelihood function with respect to the prior distribution over the parameter space, or, equivalently, using the Bayes factor, defined as the ratio of marginal likelihoods. Although the marginal likelihood is often viewed as difficult to compute, it can be estimated readily from MCMC output (Chib, 1995b; Chib and Jeliazkov, 2001) and has been used to compare censored quantile regression models (Yu and Stander, 2007; Kobayashi and Kozumi, 2012). More generally, however, the use of the marginal likelihood for Bayesian model comparison in quantile regression remains relatively limited, with notable exceptions including Maheshwari and Rahman (2023), Jeliazkov et al. (2023), and Song et al. (2026).

To estimate the marginal likelihood for a model $\mathcal{M}$, we first express it as the ratio of the likelihood function times the prior density to the posterior density,

$$m(y|\mathcal{M}) = \frac{f(y|\mathcal{M},\Theta)\,\pi(\Theta|\mathcal{M})}{\pi(\Theta|\mathcal{M},y)}, \tag{14}$$

where Θ denotes the parameter vector. Equation (14) is an identity that holds for all Θ, but is typically evaluated at a high-density point Θ^* (e.g., the posterior mean or mode) to reduce estimation variability. The likelihood $f(y|\mathcal{M},\Theta^*)$ and the prior ordinate $\pi(\Theta^*|\mathcal{M})$ are typically available, so the primary task is to estimate the posterior ordinate $\pi(\Theta^*|y,\mathcal{M})$ from MCMC output. In the QCLD model, all the conditional distributions in Algorithm 1 have known form, so the posterior ordinate is estimated using the approach of Chib (1995b).

We let $\Theta = (\beta, \sigma, \Psi)$ and factorize the posterior ordinate $\pi(\Theta^*|\mathcal{M}, y)$ as

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

where $\Theta^* = (\beta^*, \sigma^*, \Psi^*)$ and the conditioning on $\mathcal{M}$ is suppressed for simplicity. Throughout, the variables (z, α, ν) are integrated out before evaluating the posterior ordinates, which reduces the dimension of integration and improves numerical accuracy.

To simplify the notation that follows, let $\Theta_1 = (\Theta_2, \Psi)$ and $\Theta_2 = (z, \alpha, \nu)$. The first term, $\pi(\beta^*|y) = \int \pi(\beta^*|y,\Theta_1)\,\pi(\Theta_1|y)\,d\Theta_1$, is estimated by the Monte Carlo average $\widehat{\pi}(\beta^*|y) = E\{\pi(\beta^*|y,\Theta_1)\}$, where the expectation is approximated using draws from $\pi(\Theta_1|y)$ obtained in the main run. To estimate $\pi(\sigma^*|y,\beta^*) = \int \pi(\sigma^*|y,\beta^*,\Theta_2)\,\pi(\Theta_2|y,\beta^*)\,d\Theta_2$, we form a reduced run of Algorithm 1

with β fixed at β^*, simulating draws from $\pi(\Theta_2|y,\beta^*)$ that yield $\widehat{\pi}(\sigma^*|y,\beta^*) = E\{\pi(\sigma^*|y,\beta^*,\Theta_2)\}$. Finally, a second reduced run of Algorithm 1 with (β,σ) fixed at (β^*,σ^*) is used to estimate $\pi(\Psi^*|y,\beta^*,\sigma^*) = \int \pi(\Psi^*|y,\beta^*,\sigma^*,\Theta_2)\,\pi(\Theta_2|y,\beta^*,\sigma^*)\,d\Theta_2$, as $\widehat{\pi}(\Psi^*|y,\beta^*,\sigma^*) = E\{\pi(\Psi^*|y,\beta^*,\sigma^*,\Theta_2)\}$, where the expectation is with respect to $\pi(\Theta_2|y,\beta^*,\sigma^*)$.

With an estimate of the joint posterior ordinate in hand, the remaining quantities required are the prior ordinates and the likelihood of the QCLD model. The prior ordinates are readily available because the prior distributions for (β,σ,Ψ) have closed-form expressions, as specified in equation (5). The likelihood is evaluated marginally of (α,ν) as

$$f(y|\Theta^*) = \prod_{i=1}^{n}\left[\int\left\{\prod_{t=1}^{T_i} f_{AL}\left(y_{it}|\mu_{it},\sigma^*,p\right)^{I(y_{it}>0)} * F_{AL}(0|\mu_{it},\sigma^*,p)^{I(y_{it}=0)}\right\} f(\alpha_i|\Psi^*)\,d\alpha_i\right], \quad (16)$$

where $\mu_{it} = x_{it}'\beta^* + s_{it}'\alpha_i$, $I(\cdot)$ is the indicator function, and f_{AL} (F_{AL}) denotes the *pdf* (*cdf*) of the AL distribution. The likelihood (16) is estimated by Monte Carlo integration. Finally, having obtained estimates of all the quantities in equation (14), the logarithm of the marginal likelihood for the QCLD model is estimated as

$$\ln \widehat{m}(y) = \ln f(y|\Theta^*) + \ln \pi(\Theta^*) - \ln\left[\widehat{\pi}(\beta^*|y)\,\widehat{\pi}(\sigma^*|y,\beta^*)\,\widehat{\pi}(\Psi^*|y,\beta^*,\sigma^*)\right]. \quad (17)$$

Taken together, the blocked sampler, covariate effects, and marginal likelihood provide a complete framework for estimation, inference, and model comparison in the QCLD model. Appendix **??** documents its performance across a range of simulation designs. We now turn to the RFC application.

5 Results

5.1 QBLD Results

We estimate the binary model at each quantile $p \in \{0.1, 0.2, \ldots, 0.9\}$ on the full sample of 2,979 counties. Each specification is drawn for 10,000 MCMC iterations (20,000 at the lowest quantiles, where mixing is slower) with the first 20% discarded as burn-in. Figure 3 reports the results as covariate effects (i.e., the change in the predicted probability that a county receives RFC lending, averaged over the posterior draws and the sample). Markers are black-filled when the 95% posterior credible interval excludes zero.

We find that neighboring participation (AFAB) raises the probability of own-county participation across all quantiles, with effects ranging from roughly 1.5 to 3 percentage points. This is sizable given the unconditional probability of participation is 12.7%. Own-county participation in the prior month (FAB) is likewise positive. What stands out is how the two effects compare across the distribution. At the lowest quantile, the neighboring effect is the larger of the two. Meaning, for counties least likely to participate, a neighbor's recent participation matters most. Moving up the distribution, this ordering reverses. Own-county persistence catches up through the middle quantiles, rising steeply while the neighboring effect holds roughly flat near 3 percentage points. The relative importance of neighbors is therefore greatest where baseline participation is lowest. A neighbor's action is producing an informative signal, which is consistent with mechanisms related to higher participation leading to pooling and lower stigma or program familiarity. Importantly, the neighboring participation effect remains strong even with both closure measures included, indicating that the spillover is not simply an artifact of shared regional distress.

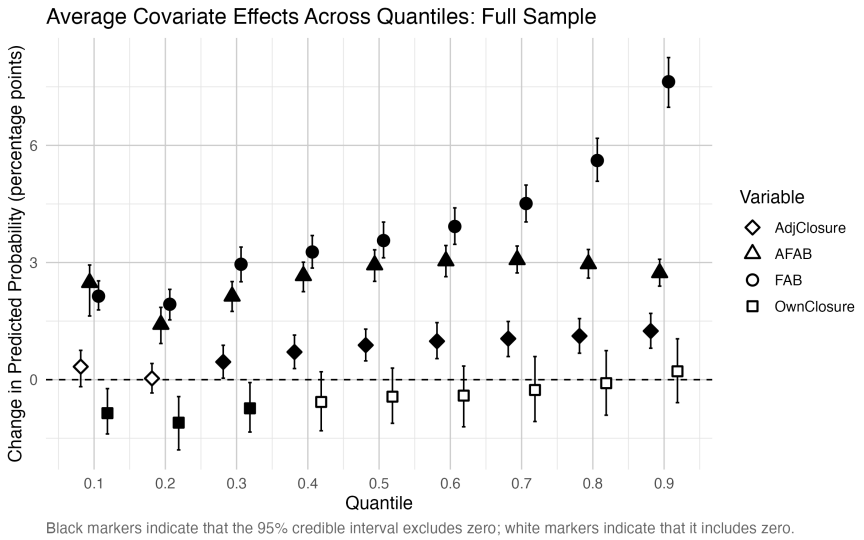


Figure 3: Average covariate effects across quantiles for the full sample. Points denote posterior mean effects in percentage points and vertical bars denote 95% credible intervals. AFAB: lagged adjacent-county RFC participation, FAB: lagged own-county RFC participation, OwnClosure: lagged own-county bank closures, AdjClosure: lagged adjacent-county bank closures.

Our adjacency measure treats all neighboring participation alike, but the two bank charter types differed in ways that speak to our mechanisms. National banks belonged to the Federal Reserve System and could borrow at the discount window, giving them an alternative source of emergency

liquidity that did not have a stigma problem (Anbil and Vossmeyer, 2019, 2021). State banks, most of which were not members, depended on the RFC as their primary channel of federal support. If neighboring participation matters because it lowers the informational and reputational barriers to seeking assistance, then it should matter most for the type of bank for which the RFC was the relevant option. To test this, we decompose our measure into adjacent state-bank participation (AFASB) and adjacent national-bank participation (AFANB) and re-estimate the model.

Figure 4 reports the results. Both channels are positive at every quantile, but the state-bank channel is consistently the larger of the two. That neighboring participation diffused most strongly among state banks is consistent with state banks facing higher familiarity and stigma costs given they most likely have not interacted with a federal facility before and could not approach an alternative, unstigmatized facility.

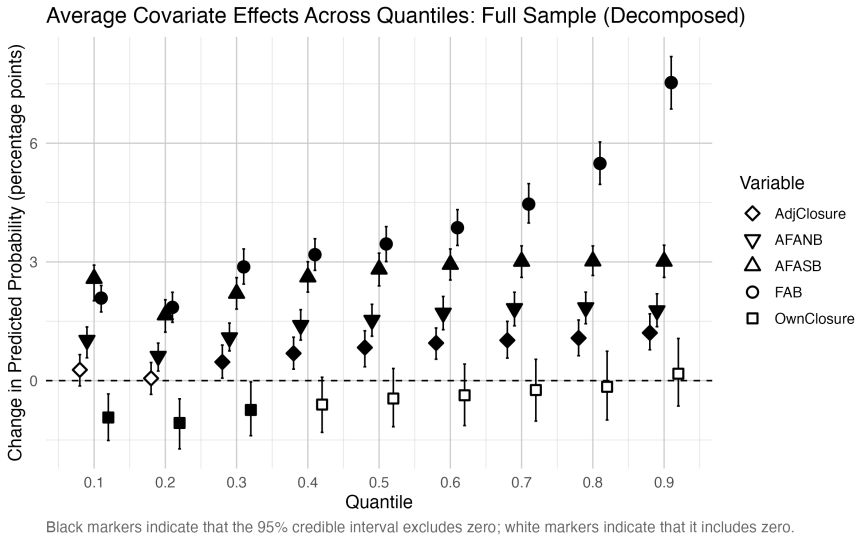


Figure 4: Average covariate effects across quantiles with adjacent-county RFC participation decomposed by bank type. AFASB and AFANB denote lagged adjacent-county participation by state-chartered and nationally chartered banks, respectively.

Two features of the panel motivate a set of subsample analyses. First, the own-county regressors are only well defined when a county has enough banks for them to vary meaningfully. We therefore restrict attention to counties with at least three banks. Second, participation behavior may differ systematically with county size where small, sparsely banked counties differ from larger ones in credit demand and the visibility of any RFC involvement. To examine this, we split the sample into the bottom quartile of counties by population and the remaining three-quarters, estimating

the model separately within each.

Figure 5 shows the results when we restrict to counties with at least 3 banks. The main results are intact. Neighboring participation (AFAB) is positive at every quantile, ranging from roughly 1.5 to 3.5 percentage points, and own-county participation (FAB) is positive throughout, climbing steeply in the upper tail. The rest-75% subsample (Figure 6b) closely mirrors these results and those of the full sample.

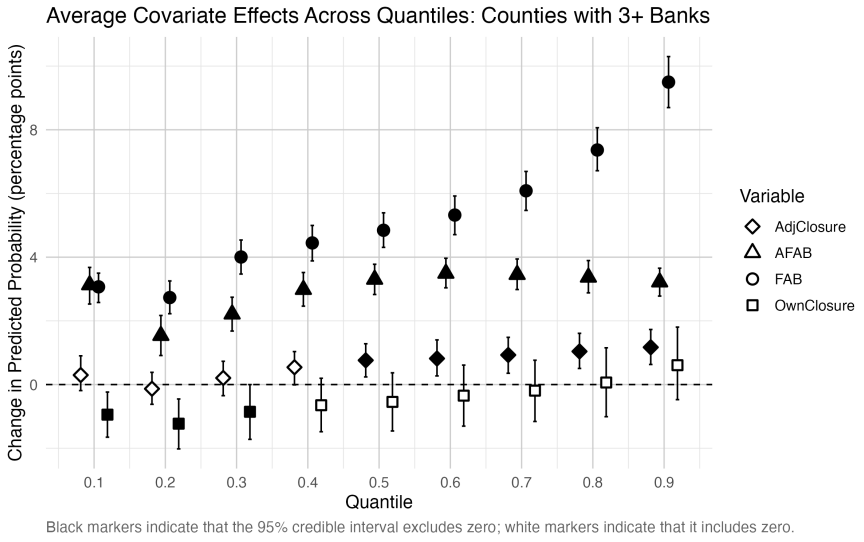
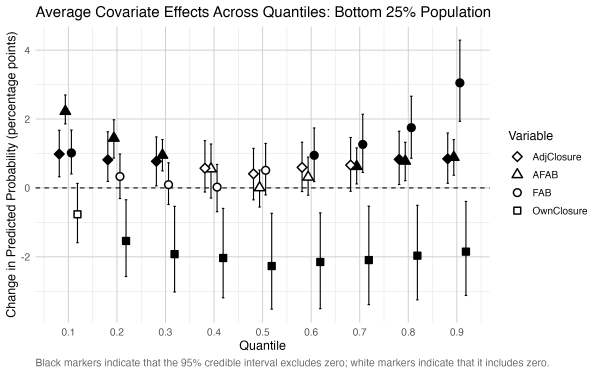


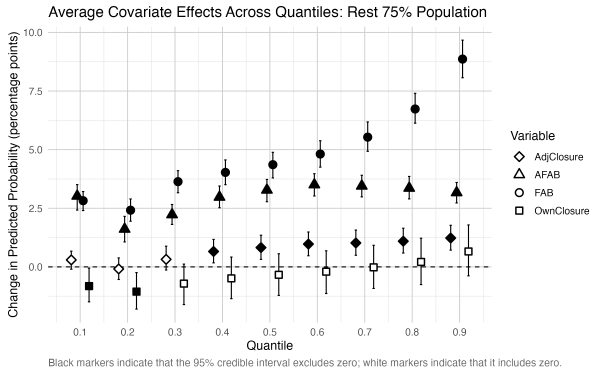
Figure 5: Average covariate effects across quantiles for counties with at least three banks.

The bottom-25% subsample (Figure 6a) tells a different and noisier story. Here the neighboring effect is statistically different from 0 and positive only at the lowest and highest quantiles. The credible intervals are markedly wider throughout, reflecting the smaller sample and the sparser banking activity in these counties. Importantly, however, the neighboring participation spillover remains a meaningful determinant even for the least-participating, least-connected counties.

Lastly, we explore the results when we restrict the sample to counties in each of the Federal Reserve districts to confirm that a particular region is not driving our results. We re-estimate the model separately within each of the twelve Federal Reserve districts, at quantiles $\tau \in \{0.1, 0.5, 0.9\}$. Figure 7 collects the results. We find that at the lowest quantile, the neighboring participation effect (AFAB) is positive and statistically different from 0 in all twelve districts, with point estimates ranging from roughly 2 to 6 percentage points. The lower-tail spillover is thus not a feature of any single region.



(a) Bottom 25 percent by population



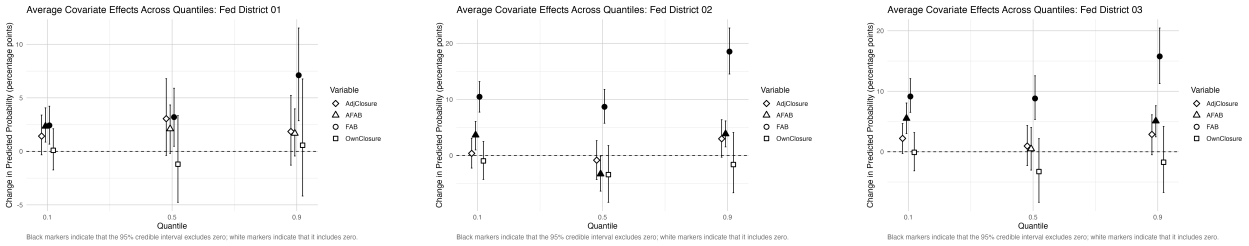
(b) Remaining 75 percent by population

Figure 6: Average covariate effects across quantiles by county population.

The binary model results are consistent across the full sample, the population and bank-density subsamples, and all twelve Federal Reserve districts. The probability of a county's RFC participation increases with the recent participation of its neighbors, by several percentage points against a base rate under thirteen percent. Because every specification controls for own-county and neighboring closures, this is not just shared regional distress. We have two key findings that speak to the mechanism. First, we find that the neighboring effect matters most, relative to a county's own history, at the lower tail, where counties are least likely to participate, and a neighbor's example is most informative. Second, we also find that state banks, which had largely no discount-window access and thus no alternative to the RFC, were a big spillover driver. Both patterns point to participation spreading by easing stigma and unfamiliarity.

5.2 QCLD Results

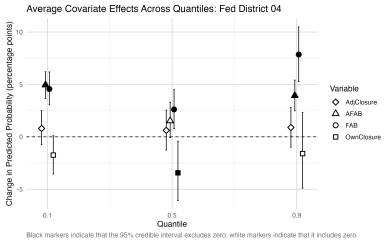
[STILL LOADING...]



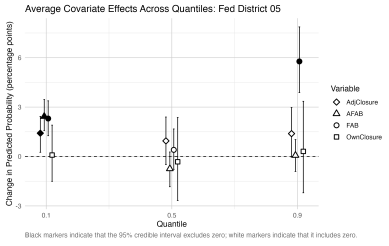
(a) Boston

(b) New York

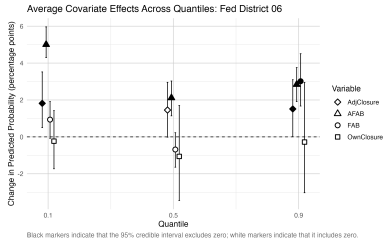
(c) Philadelphia



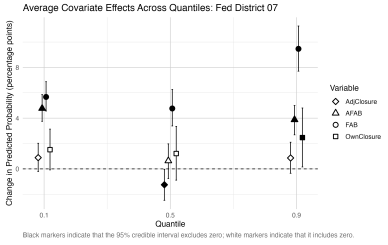
(d) Cleveland



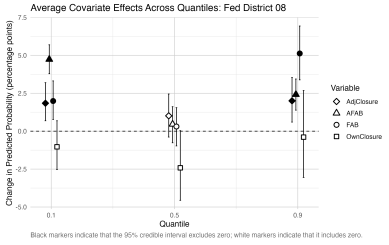
(e) Richmond



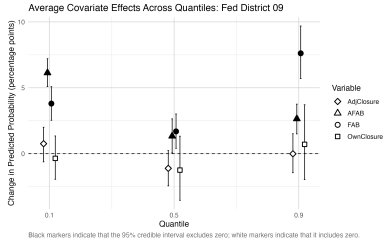
(f) Atlanta



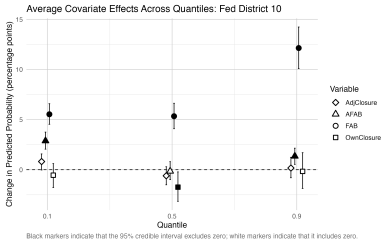
(g) Chicago



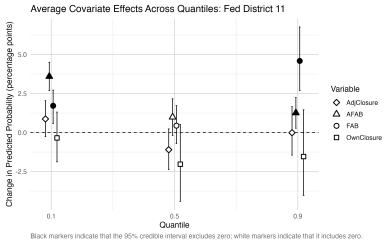
(h) St. Louis



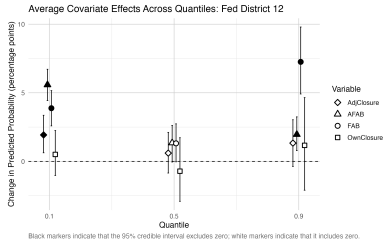
(i) Minneapolis



(j) Kansas City



(k) Dallas



(l) San Francisco

Figure 7: Average covariate effects across Federal Reserve districts. District-level models are estimated at the 0.1, 0.5, and 0.9 quantiles.

6 Conclusion

This paper studies whether participation in emergency government lending is contagious across space, that is, whether a bank's decision to seek assistance depends on whether its neighbors have recently done so. We examine this question using the Reconstruction Finance Corporation, the principal source of emergency assistance to banks during the Great Depression. We assemble a novel county-level monthly panel from digitized RFC records, and pair it with data on bank closures that let us distinguish spillovers from shared regional distress. On the methodological side, we develop a quantile regression model for censored longitudinal data, extending Bayesian quantile methods to left-censored outcomes and allowing spillover effects to vary across the full distribution of lending.

Our results indicate that participation in RFC lending exhibited positive spatial spillovers, with a county's uptake rising when its neighbors had recently participated. The effect is strongest at the lower tail of the conditional distribution (i.e., regions least likely to receive assistance), precisely where stigma and informational barriers are highest and where peer participation would send a strong signal. We also find that the spillovers are the largest for institutions with the least prior familiarity with federal credit facilities. These patterns speak to policymakers' actions during recent crises. Strategies that pool participation (e.g., Term Auction Facility) or coordinate acceptance of bank participation (e.g., TARP capital) can weaken the adverse signal of borrowing and pull in the institutions most deterred by it.

References

Amujala, S., Vossmeyer, A., and Das, S. R. (2023), "Digitization and Data Frames for Card Index Records," *Explorations in Economic History*, 87, 101469.

Anbil, S. (2018), "Managing Stigma During a Financial Crisis," *Journal of Financial Economics*, forthcoming.

Anbil, S. and Vossmeyer, A. (2019), "The Quality of Banks at Stigmatized Lending Facilities," *AEA Papers and Proceedings*, 109, 506–510.

Anbil, S. and Vossmeyer, A. (2021), "Liquidity from Two Lending Facilities," *Journal of Financial Intermediation*, 48.

Anbil, S., Carlson, M., and Styczynski, M.-F. (2023), "The Effect of the PPPLF on PPP Lending by Commercial Banks," *Journal of Financial Intermediation*, 55, 101042.

Armantier, O. and Holt, C. A. (2020), "Overcoming Discount Window Stigma: An Experimental Investigation," *The Review of Financial Studies*, 33, 5630–5659.

Armantier, O. and Holt, C. A. (2025), "Can Discount Window Stigma Be Cured? An Experimental Investigation," *The Review of Financial Studies*, Forthcoming. Earlier version: Federal Reserve Bank of New York Staff Report No. 1103, May 2024, doi:10.59576/sr.1103.

Armantier, O., Ghysels, E., Sarkar, A., and Shrader, J. (2015), "Discount Window Stigma During the 2007-2008 Finanial Crisis," *Journal of Financial Economics*, 118, 317–335.

Armantier, O., Cipriani, M., and Sarkar, A. (2024), "Discount Window Stigma After the Global Financial Crisis," Staff Report 1137, Federal Reserve Bank of New York.

Bajaj, A. (2018), "Undefeated Equilibria of the ShiTrejosWright Model Under Adverse Selection," *Journal of Economic Theory*, 176, 957–986.

Bernanke, B. (2009), "The Federal Reserve's Balance Sheet: An Update," in *Speech at the Federal Reserve Board Conference on Key Developments in Monetary Policy*.

Bertrand, M., Luttmer, E. F. P., and Mullainathan, S. (2000), "Network Effects and Welfare Cultures," *The Quarterly Journal of Economics*, 115, 1019–1055.

Beyhaghi, M. and Gerlach, J. R. (2025), "Discount Window Borrowing 2003–2019," Working paper, Available at SSRN: https://ssrn.com/abstract=5314549.

Bhargava, S. and Manoli, D. (2015), "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment," *American Economic Review*, 105, 3489–3529.

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.

Calomiris, C. W., Mason, J. R., Weidenmier, M., and Bobroff, K. (2013), "The Effects of the Reconstruction Finance Corporation Assistance on Michigan's Banks' Survival in the 1930s," *Explorations in Economic History*, 50, 525–547.

Chamberlain, G. (1984), "Panel Data," in *Handbook of Econometrics*, eds. Z. Griliches and M. D. Intriligator, vol. 2, pp. 1247–1318, Elsevier.

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

Chib, S. (1995b), "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," *Statistics and Computing*, 9, 17–26.

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.

Currie, J. (2006), "The Take-Up of Social Benefits," in *Public Policy and the Income Distribution*, eds. A. J. Auerbach, D. Card, and J. M. Quigley, pp. 80–148, Russell Sage Foundation, New York.

Dahl, G. B., Løken, K. V., and Mogstad, M. (2014), "Peer Effects in Program Participation," *American Economic Review*, 104, 2049–2074.

Ennis, H. (2018), "Interventions in Markets with Adverse Selection: Implications for Discount Window Stigma," *Journal of Money, Credit and Banking*, forthcoming.

Ennis, H. and Weinberg, J. (2013), "Over-the-counter Loans, Adverse Selection, and Stigma in the Interbank Market," *Review of Economic Dynamics*, 16, 601–616.

Finkelstein, A. and Notowidigdo, M. J. (2019), "Take-Up and Targeting: Experimental Evidence from SNAP," *The Quarterly Journal of Economics*, 134, 1505–1556.

Geithner, T. (2014), *Stress Test: Reflections on Financial Crises*, Random House.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013), *Bayesian Data Analysis*, 3rd Edition, Chapman & Hall, New York.

Geweke, J. (1991), "Efficient Simulation from the Multivariate Normal and Student-t Distributions subject to Linear Constraints," in *Computing Science and Statistics: Proceedings of the twenty-third Symposium on the Interface*, ed. E. Keramidas, pp. 571–578, Interface Foundation of North America, Inc., Fairfax.

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

Gorton, G. and Ordonez, G. (2017), "Fighting Crises with Secrecy," *Working Paper*.

Greenberg, E. (2012), *Introduction to Bayesian Econometrics*, 2nd Edition, Cambridge University Press, New York.

Jeliazkov, I. and Vossmeyer, A. (2018), "The Impact of Estimation Uncertainty on Covariate Effects in Nonlinear Models," *Statistical Papers*, 59, 1031–1042.

Jeliazkov, I., Karnawat, S., Rahman, M. A., and Vossmeyer, A. (2023), "Flexible Bayesian Quantile Analysis of Residential Rental Rates," https://doi.org/10.48550/arXiv.2305.13687.

Kobayashi, G. (2017), "Bayesian Endogenous Tobit Quantile Regression," *Bayesian Analysis*, 12, 161–191.

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

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

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.

Mason, J. R. (2001), "Do Lender of Last Resort Policies Matter? The Effects of Reconstruction Finance Corporation Assistance to Banks During the Great Depression," *Journal of Financial Services Research*, 20, 77–95.

Mason, J. R. (2003), "The Political Economy of Reconstruction Finance Corporation Assistance During the Great Depression," *Explorations in Economic History*, 40, 101–121.

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

Philippon, T. and Skreta, V. (2012), "Optimal Interventions in Markets with Adverse Selection," *American Economic Review*, 102, 1–28.

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

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

Song, J. J., Rahman, M. A., Chin, Y.-M., and Stamey, J. D. (2026), "Modeling Misclassification in Spousal Violence Reporting: Evidence from Bayesian Quantile Regression," *www.arXiv.org*, https://doi.org/10.48550/arXiv.2605.15428.

Vossmeyer, A. (2016), "Sample Selection and Treatment Effect Estimation of Lender of Last Resort Policies," *Journal of Business and Economic Statistics*, 34, 197–212.

Vossmeyer, A. (2019), "Analysis of Stigma and Bank Credit Provision," *Journal of Money, Credit and Banking*, 51, 163–194.

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

Yu, K. and Stander, J. (2007), "Bayesian Analysis of a Tobit Quantile Regression Model," *Journal of Econometrics*, 137, 260–276.

Yu, K. and Zhang, J. (2005), "A Three Parameter Asymmetric Laplace Distribution and its Extensions," *Communications in Statistics – Theory and Methods*, 34, 1867–1879.

Yuan, Y. and Yin, G. (2010), "Bayesian Quantile Regression for Longitudinal Studies with Nonignorable Missing Data," *Biometrics*, 66, 105–114.

Appendix A Non-blocked Sampling in CPQR Model

The algorithm below outlines the non-blocked MCMC sampler for the CPQR model, which is largely based on the algorithm proposed by Kobayashi and Kozumi (2012).

Algorithm 2 (Non-blocked sampling)

---

1. Let $\Lambda_i^{1/2} = D_{\tau\sqrt{\sigma\nu_i}}$. Sample $[\beta|z,\alpha,\nu,\sigma] \sim N(\tilde{\beta},\tilde{B})$, where,

$$\tilde{B}^{-1} = \left(\sum_{i=1}^{n} X_i'\Lambda_i^{-1}X_i + B_0^{-1}\right) \text{ and } \tilde{\beta} = \tilde{B}\left(\sum_{i=1}^{n} X_i'\Lambda_i^{-1}(z_i - S_i\alpha_i - \theta\nu_i) + B_0^{-1}\beta_0\right).$$

2. Sample $[\alpha_i|z,\beta,\nu,\sigma,\Psi] \sim N(\tilde{a},\tilde{A})$ for $i = 1,\ldots,n$, where,

$$\tilde{A}^{-1} = \left(S_i'\Lambda_i^{-1}S_i + \Psi^{-1}\right) \quad \text{and} \quad \tilde{a} = \tilde{A}\left(S_i'\Lambda_i^{-1}\left(z_i - X_i\beta - \theta\nu_i\right)\right).$$

3. Sample $[\nu_{it}|z_{it},\beta,\alpha_i,\sigma] \sim GIG(0.5,\tilde{\lambda}_{it},\tilde{\eta})$ for $i = 1,\ldots,n$, and $t = 1,\ldots,T_i$, where,

$$\tilde{\lambda}_{it} = \left(\frac{z_{it} - x_{it}'\beta - s_{it}'\alpha_i}{\tau\sqrt{\sigma}}\right)^2 \quad \text{and} \quad \tilde{\eta} = \left(\frac{\theta^2}{\tau^2\sigma} + \frac{2}{\sigma}\right).$$

4. Sample $[\sigma|z,\beta,\alpha,\nu] \sim IG(\tilde{n}/2,\tilde{d}/2)$, where

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

5. Sample $[\Psi|\alpha] \sim IW(\tilde{r},\tilde{R})$, where $\tilde{r} = (n + r_0)$ and $\tilde{R} = \left(\sum_{i=1}^{n}\alpha_i\alpha_i' + R_0\right)$.

6. Sample $[z_{it}|y_{it},\beta,\alpha_i,\nu_{it},\sigma]$ for all $i = 1,\ldots,n$, and $t = 1,\ldots,T_i$, from the corresponding univariate truncated normal (TN) distribution as follows,

$$[z_{it}|y_{it},\beta,\alpha_i,\nu_{it},\sigma] \quad \sim \quad \begin{cases} TN_{(-\infty,0]}\left(x_{it}'\beta + s_{it}'\alpha_i + \theta\nu_{it}, \tau^2\sigma\nu_{it}\right) & \text{if } y_{it} = 0, \\ y_{it} & \text{if } y_{it} > 0. \end{cases}$$

Appendix B The Conditional Densities for Blocked Sampling in CPQR Model

This appendix derives the conditional posterior densities for the blocked sampling algorithm in the censored panel quantile regression (CPQR) model. The key feature of the algorithm is that the regression coefficients β and latent variables z are sampled jointly in a block by integrating out the random effects α. In the first sub-step, β is sampled from a multivariate normal distribution. In the second sub-step, the censored components of each z_i are sampled from a truncated multivariate

normal distribution using the Gibbs sampler of Geweke (1991), which sequentially updates each component from its conditional univariate truncated normal distribution.

The remaining parameters are sampled using standard Gibbs updates. The random effects α_i's are sampled from an updated multivariate normal distribution, the latent weights ν are sampled element-wise from generalized inverse Gaussian (GIG) distributions, the scale parameter σ is sampled from an updated inverse-Gamma distribution, and the covariance matrix Ψ is sampled from an updated inverse-Wishart distribution.

(1). In the CPQR model: $z_i = X_i\beta + S_i\alpha_i + \theta\nu_i + D_{\tau\sqrt{\sigma\nu_i}}\,u_i$, for $i = 1, \ldots, n$, the mean and variance of z_i, obtained by integrating out α_i, are given by

$$\begin{aligned} E(z_i) &= X_i\beta + \theta\nu_i, \\ V(z_i) &= S_i\Psi S_i' + D^2_{\tau\sqrt{\sigma\nu_i}} = \Omega_i. \end{aligned}$$

First, we derive the conditional posterior of β and z_i, marginally of α_i, but conditional on other variables in the model.

1(a). Starting with β, the conditional posterior density $\pi(\beta|z, \nu, \Psi)$, marginalized over α, can be derived as,

$$\begin{aligned} \pi(\beta|z,\nu,\sigma,\Psi) &\propto \left\{ \prod_{i=1}^{n} f(z_i|\beta,\nu_i,\sigma,\Psi) \right\} \pi(\beta) \\ &\propto \exp\left\{ -\frac{1}{2}\left[\sum_{i=1}^{n} (z_i - X_i\beta - \theta\nu_i)'\Omega_i^{-1}(z_i - X_i\beta - \theta\nu_i) + (\beta-\beta_0)'B_0^{-1}(\beta-\beta_0) \right] \right\} \\ &\propto \exp\left\{ -\frac{1}{2}\left[\beta'\left(\sum_{i=1}^{n} X_i'\Omega_i^{-1}X_i + B_0^{-1} \right)\beta - \beta'\left(\sum_{i=1}^{n} X_i'\Omega_i^{-1}(z_i - \theta\nu_i) + B_0^{-1}\beta_0 \right) \right.\right. \\ &\quad \left.\left. - \left(\sum_{i=1}^{n} (z_i - \theta\nu_i)'\Omega_i^{-1}X_i + \beta_0'B_0^{-1} \right)\beta \right] \right\} \\ &\propto \exp\left\{ -\frac{1}{2}\left[\beta'\tilde{B}^{-1}\beta - \beta'\tilde{B}^{-1}\tilde{\beta} - \tilde{\beta}'\tilde{B}^{-1}\beta \right] \right\}, \end{aligned}$$

where the third line retains only the terms involving β, while the fourth line introduces,

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

Adding and subtracting $\tilde{\beta}'\tilde{B}^{-1}\tilde{\beta}$ and absorbing the term $\exp[-\frac{1}{2}\{-\tilde{\beta}'\tilde{B}^{-1}\tilde{\beta}\}]$ into the proportion-

ality constant, the square can be completed as follows,

$$\pi(\beta|z,\nu,\Psi) \propto \exp\Big\{ -\frac{1}{2}(\beta-\tilde{\beta})'\tilde{B}^{-1}(\beta-\tilde{\beta})\Big\}.$$

The resulting expression is the kernel of a multivariate normal distribution; therefore $[\beta|z,\nu,\sigma,\Psi] \sim N(\tilde{\beta},\tilde{B})$.

1(b). The conditional posterior density of the latent variable z, marginalized over α, is obtained from the joint posterior distribution (**??**) as follows:

$$\begin{aligned}\pi(z|y,\beta,\nu,\sigma,\Psi) &\propto \prod_{i=1}^{n}\Big\{\pi(z_i|\beta,\nu_i,\Psi,y_i)\Big\}\\ &\propto \prod_{i=1}^{n}\Big\{\prod_{t=1}^{T_i}\Big[I(z_{it}=y_{it})I(y_{it}>0)+I(z_{it}\le 0)I(y_{it}=0)\Big]\\ &\quad\times\exp\Big[-\frac{1}{2}(z_i-X_i\beta-\theta\nu_i)'\Omega_i^{-1}(z_i-X_i\beta-\theta\nu_i)\Big]\Big\}.\end{aligned}$$

The kernel inside the curly braces corresponds to a truncated multivariate normal distribution. Therefore,

$$[z_i|y_i,\beta,\nu_i,\sigma,\Psi] \sim TMVN_{B_i}(X_i\beta+\theta\nu_i,\Omega_i), \qquad \forall\, i=1,\ldots,n.$$

where, $TMVN$ denotes a truncated multivariate normal distribution and B_i denotes the truncation region, given by $B_i = (B_{i1}\times B_{i2}\times\cdots\times B_{iT_i})$. For $t=1,\ldots,T_i$, the set B_{it} reduces to the singleton $\{y_{it}\}$ when $y_{it}>0$, and is equal to the interval $(-\infty,0]$ when $y_{it}=0$. Consequently, the latent variables corresponding to uncensored observations are fixed at their observed values and therefore require no sampling, whereas those associated with censored observations require sampling subject to $z_{it}\le 0$.

Define two index sets: $O_i=\{t: y_{it}>0\}$ and $C_i=\{t: y_{it}=0\}$ and partition the latent variable z_i, mean vector, and covariance matrix, according to the observed and censored components:

$$z_i=\begin{pmatrix} z_i^c\\ z_i^o\end{pmatrix}, \qquad \mu_i = X_i\beta+\theta\nu_i=\begin{pmatrix}\mu_i^c\\ \mu_i^o\end{pmatrix}, \qquad \Omega_i=\begin{pmatrix}\Omega_i^{cc} & \Omega_i^{co}\\ \Omega_i^{oc} & \Omega_i^{oo}\end{pmatrix},$$

where the superscripts o and c denote the uncensored (observed) and censored components, respectively. Since $z_i^o = y_i^o$ is observed exactly, only the censored component z_i^c requires sampling. The conditional distribution of z_i^c given z_i^o is,

$$[z_i^c|z_i^o,\beta,\nu_i,\sigma,\Psi] \sim TMVN_{(-\infty,0]^{|C_i|}}\Big(\eta_i^c,\,\Sigma_i^c\Big),$$

where, the conditional mean and conditional variances are,

$$\begin{aligned}\eta_i^c &= \mu_i^c + \Omega_i^{co}(\Omega_i^{oo})^{-1}(z_i^o - \mu_i^o)\\ \Sigma_i^c &= \Omega_i^{cc} - \Omega_i^{co}(\Omega_i^{oo})^{-1}\Omega_i^{oc},\end{aligned}$$

respectively. In addition, $TMVN_{(-\infty,0]^{|C_i|}}(\cdot,\cdot)$ denotes the multivariate normal distribution truncated to the region $(-\infty,0]^{|C_i|}$, where $|C_i|$ is the number of censored variables in each z_i for $i = 1,\ldots,n$.

To sample the latent variables corresponding to censored observations, we employ the Gibbs sampler of Geweke (1991), in which each censored latent variable is drawn from its conditional univariate truncated normal (TN) distribution. Specifically, let z_{it}^c denote the t-th censored latent variable, where $t = 1,\ldots,|C_i|$. Then each censored latent variable is sampled from the conditional posterior distribution as,

$$[z_{it}^c | z_{i,-t}^c, z_i^o, \beta, \nu_i, \sigma, \Psi] \sim TN_{(-\infty,0]}\left(\eta_{it|-t}^c, \Sigma_{it|-t}^c\right),$$

where $z_{i,-t}^c$ denotes the vector of remaining censored latent variables. The conditional mean and conditional variance (both scalars) are defined as,

$$\begin{aligned}\eta_{it|-t}^c &= \eta_{it}^c + \Sigma_{t,-t}^c\left(\Sigma_{-t,-t}^c\right)^{-1}\left(z_{i,-t}^c - \eta_{i,-t}^c\right),\\ \Sigma_{it|-t}^c &= \Sigma_{t,t}^c - \Sigma_{t,-t}^c\left(\Sigma_{-t,-t}^c\right)^{-1}\Sigma_{-t,t}^c,\end{aligned}$$

respectively. Here, $\eta_i^c = X_i^c\beta + \theta\nu_i^c$, and $\eta_{i,-t}^c$ is obtained by removing the t-th element from η_i^c. Likewise, $\Sigma_{t,t}^c$ denotes the (t,t)-th element of Σ_i^c, while $\Sigma_{t,-t}^c$ and $\Sigma_{-t,-t}^c$ denote the corresponding row and principal submatrix, respectively. The censored latent variables are updated sequentially until all elements of z_i^c have been sampled.

(2). The conditional posterior density of the random effects parameters α_i for $i = 1,\ldots,n$ is obtained from the joint posterior distribution (**??**) as follows,

$$\begin{aligned}&\pi(\alpha_i|z_i,\beta,\nu_i,\sigma,\Psi) \propto f(z_i|\beta,\alpha_i,\nu_i,\sigma)\,\pi(\alpha_i|\Psi)\\ &\propto \exp\left\{-\frac{1}{2}\left[(z_i - X_i\beta - S_i\alpha_i - \theta\nu_i)'\Lambda_i^{-1}(z_i - X_i\beta - S_i\alpha_i - \theta\nu_i) + \alpha_i'\Psi^{-1}\alpha_i\right]\right\}\\ &\propto \exp\left\{-\frac{1}{2}\left[\alpha_i'\Big(S_i'\Lambda^{-1}S_i + \Psi^{-1}\Big)\alpha_i - \alpha_i'\Big(S_i'\Lambda_i^{-1}\big(z_i - X_i\beta - \theta\nu_i\big)\Big)\right.\right.\\ &\qquad\left.\left. - \Big(\big(z_i - X_i\beta - \theta\nu_i\big)'\Lambda_i^{-1}S_i\Big)\alpha_i\right]\right\}\\ &\propto \exp\left\{-\frac{1}{2}(\alpha_i - \tilde{a})'\tilde{A}^{-1}(\alpha_i - \tilde{a})\right\},\end{aligned}$$

where $\Lambda_i^{1/2} = \text{diag}(\tau\sqrt{\sigma\nu_{i1}}, \cdots, \tau\sqrt{\sigma\nu_{iT_i}})$. The third line omits all terms that do not involve α_i, and the fourth line completes the square, where

$$\tilde{A}^{-1} = \left(S_i'\Lambda_i^{-1}S_i + \Psi^{-1}\right) \quad \text{and} \quad \tilde{a} = \tilde{A}\left(S_i'\Lambda_i^{-1}\big(z_i - X_i\beta - \theta\nu_i\big)\right),$$

are the posterior precision and posterior mean, respectively. The result is a kernel of a normal distribution, hence, $[\alpha_i|z_i, \beta, \nu_i, \sigma, \Psi] \sim N(\tilde{a}, \tilde{A})$ for $i = 1, \ldots, n$.

(3). The conditional posterior density of ν_{it} is obtained from the joint posterior distribution (**??**) by collecting terms involving ν_{it}. Since the elements of ν are conditionally independent, each ν_{it} is updated separately as,

$$\begin{aligned}
\pi(\nu_{it}|z_{it}, \beta, \alpha_i, \sigma) &\propto \left(2\pi\tau^2\sigma\nu_{it}\right)^{-1/2} \exp\left\{ -\frac{1}{2\tau^2\sigma\nu_{it}}\big(z_{it} - x_{it}'\beta - s_{it}'\alpha_i - \theta\nu_{it}\big)^2 - \frac{\nu_{it}}{\sigma}\right\} \\
&\propto \nu_{it}^{-1/2} \exp\left\{ -\frac{1}{2}\left[\left(\frac{z_{it} - x_{it}'\beta - s_{it}'\alpha_i}{\tau\sqrt{\sigma}}\right)^2 \nu_{it}^{-1} + \left(\frac{\theta^2}{\tau^2\sigma} + \frac{2}{\sigma}\right)\nu_{it}\right]\right\} \\
&\propto \nu_{it}^{-1/2} \exp\left\{ -\frac{1}{2}\Big(\tilde{\lambda}_{it}\nu_{it}^{-1} + \tilde{\eta}\nu_{it}\Big)\right\},
\end{aligned}$$

where the second line omits all terms that do not depend on ν_{it}, and the third line introduces the terms,

$$\tilde{\lambda}_{it} = \left(\frac{z_{it} - x_{it}'\beta - s_{it}'\alpha_i}{\tau\sqrt{\sigma}}\right)^2 \quad \text{and} \quad \tilde{\eta} = \left(\frac{\theta^2}{\tau^2\sigma} + \frac{2}{\sigma}\right).$$

This expression is recognized as the kernel of a generalized inverse Gaussian (GIG) distribution. Hence, we have $[\nu_{it}|z_{it}, \beta, \alpha_i, \sigma] \sim GIG(0.5, \tilde{\lambda}_{it}, \tilde{\eta})$ for $t = 1, \ldots, T_i$ and $i = 1, \ldots, n$.

(4). The conditional posterior density of σ is obtained from the joint posterior distribution (**??**) by retaining all terms involving σ. This is shown below.

$$\begin{aligned}
\pi(\sigma|z, \beta, \alpha, \nu) &\propto \sigma^{-\left(\sum_{i=1}^n T_i\right)/2} \exp\left\{ -\frac{1}{2\tau^2\sigma\nu_{it}} \sum_{i=1}^{n}\sum_{t=1}^{T_i} \big(z_{it} - x_{it}'\beta - s_{it}'\alpha_i - \theta\nu_{it}\big)^2 \right\} \\
&\times \sigma^{-\left(\sum_{i=1}^n T_i\right)} \exp\left\{ -\frac{1}{\sigma}\sum_{i=1}^{n}\sum_{t=1}^{T_i}\nu_{it}\right\} \times \sigma^{-\left(\frac{n_0}{2}+1\right)} \exp\left\{ -\frac{d_0}{2\sigma}\right\} \\
&\propto \sigma^{-\left(n_0 + 3\sum_{i=1}^n T_i\right)/2 - 1} \exp\left\{ -\frac{1}{2\sigma}\left[d_0 + 2\sum_{i=1}^{n}\sum_{t=1}^{T_i}\nu_{it} + \sum_{i=1}^{n}\sum_{t=1}^{T_i}\frac{\left(z_{it} - x_{it}'\beta - s_{it}'\alpha_i - \theta\nu_{it}\right)^2}{\tau^2\nu_{it}}\right]\right\} \\
&\propto \sigma^{-\left(\frac{\tilde{n}}{2}+1\right)} \exp\left\{ -\frac{\tilde{d}}{2\sigma}\right\},
\end{aligned}$$

which is recognized as the kernel of an inverse-Gamma distribution, where $[\sigma|z, \beta, \alpha, \nu] \sim IG(\tilde{n}/2, \tilde{d}/2)$

and the terms within parenthesis are,

$$\tilde{n} = n_0 + 3\sum_{i=1}^{n} T_i \quad \text{and} \quad \tilde{d} = d_0 + 2\sum_{i=1}^{n}\sum_{t=1}^{T_i} \nu_{it} + \sum_{i=1}^{n}\sum_{t=1}^{T_i} \frac{(z_{it} - x_{it}'\beta - s_{it}'\alpha_i - \theta\nu_{it})^2}{\tau^2\nu_{it}}.$$

(5). Finally, retaining all terms involving Ψ in (**??**) yields the conditional posterior density, as derived below,

$$\begin{aligned}
\pi(\Psi|y,\alpha) &\propto |\Psi|^{-\frac{n}{2}} \exp\left\{ -\frac{1}{2}\sum_{i=1}^{n} \alpha_i'\Psi^{-1}\alpha_i \right\} \times |\Psi|^{-\left(\frac{r_0+l+1}{2}\right)} \exp\left\{ -\frac{1}{2}\operatorname{tr}(\Psi^{-1}R_0) \right\} \\
&\propto |\Psi|^{-\left(\frac{n+r_0+l+1}{2}\right)} \exp\left\{ -\frac{1}{2}\operatorname{tr}\left[\Psi^{-1}\left(\sum_{i=1}^{n}\alpha_i\alpha_i' + R_0\right)\right]\right\} \\
&\propto |\Psi|^{-\left(\frac{\tilde{r}+l+1}{2}\right)} \exp\left\{ -\frac{1}{2}\operatorname{tr}\left[\Psi^{-1}\tilde{R}\right]\right\}.
\end{aligned}$$

This is recognized as the kernel of an inverse-Wishart distribution, whereby $[\Psi|y,\alpha] \sim IW(\tilde{r},\tilde{R})$ with $\tilde{r} = n + r_0$ and $\tilde{R} = (\sum_{i=1}^{n}\alpha_i\alpha_i' + R_0)$.