Federal Reserve Bank of Cleveland Working Paper Series

Tail Sensitivity of US Bank Net Interest Margins: A Bayesian Penalized Quantile Regression Approach

Nicholas Fritsch

Working Paper No. 25-09

March 2025

Suggested citation: Fritsch, Nicholas. 2025. "Tail Sensitivity of US Bank Net Interest Margins: A Bayesian Penalized Quantile Regression Approach." Working Paper No. 25-09. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202509.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Tail Sensitivity of US Bank Net Interest Margins: A Bayesian Penalized Quantile Regression Approach

Nicholas Fritsch*†

2025-02-24

Abstract

Bank net interest margins (NIM) have been historically stable in the US on average, but this stability deteriorated in the post-2020 period, particularly in the tails of the distribution. Recent literature disagrees on the extent to which banks hedge interest rate risk, and past literature shows that credit risk and persistence are also important considerations for bank NIM. I use a novel approach to Bayesian dynamic panel quantile regression to document heterogeneity in US bank NIM estimated sensitivities to interest rates, credit risk, and own persistence. I find increased sensitivity to interest rates in the tails of the conditional NIM distribution during the post-2020 period, driven by increased interest rate sensitivities of bank loans and deposits. Density forecast evaluation shows that the model forecasts outperform frequentist benchmark models, and standard tail risk measures show that risks to bank NIM have material implications for bottom-line measures of bank profitability.

Keywords: Net Interest Margins, Interest Rate Risk, Bayesian Quantile Regression, Dynamic Panel, Density Forecasting

JEL Codes: C21, C23, E43, G21

*Federal Reserve Bank of Cleveland, nicholas.fritsch@clev.frb.org.

†I am grateful to Gary Koop and Stuart McIntyre for excellent guidance, comments and suggestions. I also thank Paola Boel, Todd Clark, Ben Craig, James Mitchell and Ned Prescott for helpful comments. The views stated herein are those of the author and are not necessarily those of the Federal Reserve Bank of Cleveland or of the Board of Governors of the Federal Reserve System.

1 Introduction

Net interest margins (NIM) of US banks have been historically stable throughout economic fluctuations, but this stability appears to have been broken in the post-pandemic period. In the second quarter of 2021, the FDIC noted that average commercial bank NIM had reached a record low as a result of declines in interest income outpacing declines in interest expenses.[1] This has contributed to the increased attention to the interest rate risk of the US banking sector for academics and bank regulators in the wake of the March 2023 US banking turmoil. Historically, depressed bank NIMs have led to failures in the banking sector. For example, during the US savings and loans (S&L) crisis in the 1980s, short-term interest rates rose above long-term interest rates, resulting in interest expenses rising above interest incomes for certain institutions, and ultimately a substantial amount of losses and bank failures in the S&L banking sector. Recent history has also shown that individual outlier banks, such as Silicon Valley Bank, can impose negative impacts on the banking sector without holding a disproportionately large amount of total assets.

Recent academic work has focused on interest rate risk in two primary components of banks' balance sheets: (1) the impact on unrealized losses of bank fixed-income portfolios (for example, Flannery and Sorescu (2023), Jiang et al. (2023)), and (2) the impact on interest-rate-sensitive cash flows captured by the NIM of banks, representing profits that are the difference between interest income and interest expense, scaled by a bank's interest-earning assets (for example, Abdymomunov, Gerlach, and Sakurai (2023)). This paper focuses on the latter. There exist differing views in the literature on the ability of banks to hedge macroeconomic interest rate impacts on bank NIMs (for example, Drechsler, Savov, and Schnabl (2021), Williams (2020), Begenau and Stafford (2022), Abdymomunov, Gerlach, and Sakurai (2023)). Additionally, past work has documented the importance of persistence and macroeconomic credit risk in explaining bank net interest margins, which the above studies do not explicitly address in their empirical frameworks. I ask the following questions: Is there heterogeneity in sensitivity to interest rates, credit risk, and persistence across the conditional NIM distribution, and does estimation of this sensitivity improve predicted outcomes? It is

[1] See https://www.fdic.gov/analysis/quarterly-banking-profile/qbp/2021jun/qbp.pdf.

useful for regulators and bank supervisors to better understand and predict the conditional distribution of bank NIMs, particularly when concerned about tail risks.

I contribute to this literature by documenting heterogeneity in sensitivities to macroeconomic interest rates, credit risk factors, and persistence of bank-level NIMs in a quantile regression framework, and evaluating the resulting density forecasts against benchmark models using proper scoring rules of the density predictions. Covas, Rump, and Zakrajšek (2014) find that a dynamic panel quantile regression framework generally outperforms more commonly used linear models when forecasting components of bank capital conditional on macroeconomic outcomes, including NIM, given the models' ability to capture nonlinearities. Similar to Covas, Rump, and Zakrajšek (2014), I estimate a dynamic panel quantile regression framework. However, I focus exclusively on bank NIMs and their underlying interest income and expense components, I include a much larger sample of banks estimated over a relatively longer time span, and I use a novel Bayesian approach to panel quantile regression as proposed by Aghamohammadi and Mohammadi (2017). Shrinkage in the Bayesian setting can lead to improved estimation and forecasting outcomes. The adaptive LASSO shrinkage used in this framework addresses the incidental parameters problem in fixed-effect nonlinear estimation and the potential bias associated with using a single LASSO shrinkage parameter, while allowing for quantile-varying individual effects and having the added benefit of enjoying the desirable "oracle" properties, behaving as if the true underlying model were known (Zou (2006)). These points of departure allow for a more flexible approach relative to the framework used by Covas, Rump, and Zakrajšek (2014), who use a single LASSO shrinkage parameter and a quantile-invariant individual effect for a small sample of large banks, and for more accurate out-of-sample model forecasts during the 2020-2022 period.

I estimate the model on quarterly regulatory filing data from a sample of 313 US bank holding companies (BHC) during 1998-2022. I document heterogeneity in the conditional quantile estimates of interest rate, credit risk and persistence variables commonly used in the literature to model NIM across the conditional distribution. I find that a 100 basis point increase in the 3-month Treasury bill (Treasury spread) is associated with 5, 3, and 8 (7, 5, and 9) basis point increases in NIM at the 5th, 50th and 95th conditional quantiles, respectively. A 100 basis point increase in the credit risk spread is associated with a 4 basis point decrease in NIM at the 5th conditional quantile and a 4 basis point increase in NIM at the 95th conditional quantile. Time period subsample analysis finds that banks' NIM sensitivity to interest rates has increased since 2020 and this increase is greater in the conditional tails of the distribution, driven by higher interest rate sensitivities for loan interest income and demand deposit interest expense, and is relatively more pronounced for the largest banks. Importantly, these increased tail sensitivities are economically meaningful for banks when interest rate or credit risk movements are large.

I use the conditional quantile estimates to produce full density forecasts for each bank in the sample. In a forecasting exercise, I examine whether conditional density estimates from my primary model specification outperform those from fixed-effect OLS, pooled quantile regression, and the fixed-effect quantile autoregressive model framework of Covas, Rump, and Zakrajšek (2014). I find that the Bayesian panel quantile regression framework statistically outperforms several competing benchmark models in forecasting pseudo out-of-sample one-period ahead outcomes over the post-pandemic period. These results are suggestive of heterogeneity in parameter estimates and Bayesian shrinkage in the individual effects leading to improvement in the prediction of bank NIMs. I also find that a model specification that includes macroeconomic interest rate and credit risk factors outperforms a model that excludes these macroeconomic covariates during the out-of-sample period, suggesting that distributional sensitivity to interest rates, which were relatively volatile in the post-2020 period, matter for bank NIMs. Additional checks show that the model results are robust to informative prior selection and forecasts that allow for additional variation across the size distribution of banks.

Of use to regulators and bank examiners, the individual bank conditional density forecasts of NIMs allow for evaluations of probabilistic risks. Examining the individual banks' density forecasts, outlier banks with relatively larger probability mass at near-zero levels of NIM are observed in the 2020-2021 period, while bank densities had largely improved by the end of 2022. In an aggregate setting, the Bayesian framework that I use appears to capture relatively larger upper and lower tail risks across aggregated density forecasts of bank NIMs relative to competing model frameworks, while still placing the highest probability closest to the aggregated actual outcomes. I compute the expected shortfall across the individual bank

NIM density forecasts and show that the size of the upside and downside risks are material across the distribution relative to bottom-line bank profitability, return on assets (ROA), where the shortfall is sizable enough to move a portion of the conditional distribution to negative ROA levels.

The remainder of the paper proceeds as follows. Section 2 reviews the relevant literature. Section 3 provides background information on recent developments of bank NIM, and documented determinants of NIM in the theoretical and empirical literature. Section 4 describes the empirical framework used in the analysis. Section 5 describes the data used for the analysis. Section 6 provides the empirical results. Section 7 examines forecasting results of the model. Section 8 provides robustness checks of the results. Section 9 looks at tail risk implications of the density forecasts, and section 10 concludes.

2 Literature

Three primary strands of the literature are related to this paper. The first is bank net interest margin sensitivities to macroeconomic interest rates, credit risk, and own persistence. Theoretical and empirical work has found a positive relationship between bank NIM and interest rate factors (for example, Alessandri and Nelson (2015); Borio, Gambacorta, and Hofmann (2017); Covas, Rump, and Zakrajšek (2014); English, Van Den Heuvel, and Zakrajšek (2018); Hirtle et al. (2016)). The empirical literature that examines interest rate impacts on NIM typically relies on the impact of factors that summarize the yield curve, namely, factors that capture the level and slope. The level is typically measured by short-term risk-free interest rates and the slope is measured by a spread between relatively long- and short-term risk-free rates (for example, Borio, Gambacorta, and Hofmann (2017); Alessandri and Nelson (2015); Hirtle et al. (2016); Abdymomunov, Gerlach, and Sakurai (2023)). The level of the yield curve captures expenses associated with the deposit franchise, which banks can profit from by pricing deposits below market rates. If banks can adjust liability expense rates more slowly than interest rates, and loans originate under the new interest rate environment, then an increase in interest rates should result in a larger increase in interest income relative to interest expense, thus increasing bank NIMs (Borio, Gambacorta, and Hofmann (2017);

Alessandri and Nelson (2015)). Alessandri and Nelson (2015) further note that banks' ability to adjust liability expenses may be constrained by the zero lower bound, suggesting a an important limitation of this tool in low interest rate environments. The slope of the yield curve captures the returns to maturity transformation, so that an increase in the slope of the yield curve is expected to increase bank net interest margins (Alessandri and Nelson (2015); Borio, Gambacorta, and Hofmann (2017), English, Van Den Heuvel, and Zakrajšek (2018), Samuelson (1945)).

Theoretical and empirical work shows that an additional concern is the impact of credit spreads, so that factors capturing macroeconomic and firm-level credit risk are also relevant for bank interest pricing decisions. Credit risk matters for bank NIMs because increased macroeconomic credit risk should create incentives for banks to increase loan risk pricing to offset potential losses (Angbazo (1997)). Other literature notes that, while this may be true in the long run, the short run implications of increased credit risk may include banks shifting their lending toward less risky and lower-yielding assets; thus, interest income generated from lending decreases (Hanweck and Ryu (2005), Zarruk and Madura (1992)).

The academic literature has also documented the importance of own persistence of bank net interest margins (English (2002), Flannery (1981)). This is partly a result of bank assets and liabilities that span multiple maturity horizons, which presents difficulty in immediately repricing the balance sheet at prevailing market rates. From a forecasting perspective, Covas, Rump, and Zakrajšek (2014) note the importance of the autoregressive terms included in their empirical model for bank profits in capturing changes in the conditional distribution during extended periods of income losses. This inclusion of persistence is also used in practical application, as models of bank revenues used by the Federal Reserve for its stress-testing purposes are specified using four lags of the dependent variable as of the most recent publication of methodology.[2]

The second relevant strand of the literature is related to the ability of banks to successfully hedge interest rate risk, and the extent to which interest rate risk is heterogeneous across banks. Flannery (1981) finds that large banks hedge interest rate risk by matching average

[2]See https://www.federalreserve.gov/publications/files/2024-march-supervisory-stress-test-methodology.pdf.

maturities in their asset and liability portfolios. English (2002) concludes that bank selection of assets, liabilities, deposit rates, and hedging activities explains in part the stability of observable NIMs as part of an international study. Drechsler, Savov, and Schnabl (2021) examine banks' ability to hedge interest rate risk arising from maturity transformation as a result of market power over the deposit franchise, such that asset and liability sensitivities to interest rates are matched and NIMs are stable across various interest rate environments. Williams (2020) comes to conclusions similar to those of as Drechsler, Savov, and Schnabl (2021) for the average bank but emphasizes that there is heterogeneity in the ability of banks to hedge. Begenau and Stafford (2022) find that matching interest income and expense sensitivities to interest rates are a consequence, but not the causal mechanism, that leads to stable net interest margins. Abdymomunov, Gerlach, and Sakurai (2023) conclude that banks, on average and in aggregate, had increased exposure to interest rate risk in the 2020-2021 pandemic period, and that this exposure was greater for small banks relative to large banks. There is additional recent evidence that banks' net interest margin sensitivity to interest rates is heterogeneous across bank size, balance sheet composition, and time periods (Sengupta and Xue (2022); Laliberte and Sengupta (2024)). In this paper, I rely on quantile regression as a natural framework for dealing with the observed and unobserved bank heterogeneity across the conditional distribution.

The third strand of the relevant literature is the stress testing literature that includes models of bank net interest margins. Most related to this paper, Covas, Rump, and Zakrajšek (2014) use a dynamic panel quantile regression framework estimated on a sample of the 15 largest US banks that would be subject to the Federal Reserve's Dodd-Frank Act stress testing requirements (DFAST) over the period 1997-2011. They show that conditional density estimation of the components of bank capital using a quantile regression framework offers improved out-of-sample forecasting performance, owing to the ability of quantile regression to capture nonlinearities present when observing losses in stressful environments. While they do not detail the net interest margin component of their satellite model approach to bank capital, they do show that some interest rate sensitivities appear relatively constant across the conditional distribution. Giglio et al. (2021) employ a similar quantile regression framework to examine forecasts of net trading income for 54 European Union banks subject to EU stress testing requirements from 2016-2020 using a method-of-moments panel data quantile regression framework, and adopt the approach of Adrian, Boyarchenko, and Giannone (2019) in constructing continuous density forecasts by fitting a parsimonious set of quantile forecasts to a skewed t-distribution to estimate measures of tail risk in the cross-section of banks. Unlike in these studies, I focus exclusively on bank net interest margins for a relatively larger merger-adjusted sample of 313 US bank holding companies over a longer time period spanning 1998-2022, using a Bayesian approach to panel data quantile regression.

3 Bank Net Interest Margins

NIM is a measure of bank profits that are a function of prevailing interest rates, capturing bank profits from the interest rate spread earned on maturity transformation inherent in the traditional "borrow short and lend long" business model of typical banking institutions. NIM is defined as the difference between the interest income and interest expense generated by an institution, normalized by the interest-earning assets of the institution. For institution i at time t, NIM is computed as

$$NIM_{it} = \frac{Net\ Interest\ Income_{it}}{Interest\ Earning\ Assets_{it}} = \frac{Interest\ Income_{it} - Interest\ Expense_{it}}{Interest\ Earning\ Assets_{it}}$$

Revenues captured by NIM are economically meaningful for banks. For example, the interest income component of NIM on average accounts for roughly 80 percent of the total income for US bank holding companies (henceforth "banks").

Figure 1 shows the average NIM for banks from 1998-2022. A long-term downward trend is visible over the period shown. This has been explained in part by a general decrease in the level of interest rates during that period (Di Lucido, Kovner, and Zeller (2017)). Prior to 2020, the figure illustrates that bank NIM is historically stable across economic fluctuations on average. However, a break in the historical stability is apparent beginning in 2020. From 2018 Q4 to 2022 Q1, average NIM decreased from 3.54 to 2.85 percent of interest-earning assets, a 19 percent decline and a historically low level of average NIM. Once the 2022 monetary policy tightening cycle began, average NIM recovered quickly toward pre-pandemic levels, increasing to 3.48 percent of interest earning assets by 2022 Q4.

The stability of bank NIM is apparent when compared to fluctuations in macroeconomic interest rates and credit risk. Figure 2 shows average NIM for banks plotted against the 3-month Treasury bill and 10-year Treasury rates from 1998-2022. The correlation in the long-term decline between the levels of average bank NIM and interest rates is visible. Also apparent from this figure is the fact that bank NIM remains relatively stable prior to 2020 against the backdrop of relatively more volatile interest rates. From 1998-2019, the standard deviation of bank NIM was 0.33, while the standard deviation of the 3-month Treasury bill rate was 1.94. Figure 3 shows average NIM for banks plotted against the spread between a measure of macroeconomic credit risk, a BBB corporate bond reference rate, and the 10-year Treasury bond rate, from 1998-2022. The volatility in the credit spread is most pronounced during the 2007-2009 great financial crisis (GFC), yet bank NIM declines by a relatively small amount during the same period.

Figure 4 shows the cross-sectional distribution of bank NIM in terms of the median, interquartile range, and 5th-95th percentiles. Over time, there is evidence of skewness in the upper and lower tails of the distribution. Lower tail skewness is pronounced beginning in 2019 Q2, after which the distance between the median and lower 5th percentile is between 45 to 88 basis points greater than the distance between the median and the upper 5th percentile, coinciding with a period of declining short-term interest rates until Q1 2022. Similarly in the GFC period, the distance between the median and lower 5th percentile was 60 basis points greater than the distance between the median and the upper 5th percentile in 2009 Q4. These summary statistics suggest the possibility of distributional heterogeneity in the stability of the distribution across economic fluctuations.

Decomposing bank NIM into interest income and interest expense shares of interest-earning assets, Figure 5 plots the distributions for US banks. The stability of NIM is typically attributed to the ability of banks to maintain a relatively constant spread between interest income and expense. However, the figure illustrates that the ability of banks to adjust deposit pricing or the maturity structure of the liability portfolio is constrained in periods when both interest expenses are already near the zero-lower bound. This is visible in 2015, when a period of low interest rates caused the distribution of interest expenses to compress to near zero while the interest income distribution has a relatively larger spread about the median. This is particularly apparent in 2021 Q4, when the lower 5th percentile of the bank interest income distribution dropped to a historical low of 1.39 percent of interest-earning assets but interest expenses remained near zero levels, having not increased enough prior to the pandemic period to allow for further decreases, so that bank adjustment options were already limited when short-term interest rates quickly moved to near-zero levels.

Figure 6 further decomposes interest income and expense as a percent of interest-earning assets into their primary constituents. On the income side, I decompose interest income into cash flows from loans, securities, and all other sources of interest income. On the liability side, I decompose interest expenses into demand deposits, time deposits, and all other sources of interest expense. These plots demonstrate that the primary sources of income and expenses historically have been loans, securities, and deposits. However, in the post-GFC period, income from securities and expenses for deposits have migrated to historically low levels, often near the zero lower bound.

The denominator of NIM, interest-earning assets, represents roughly 90 percent of total bank assets on average from 1998-2022. Interest earning assets are computed as the sum of reported quarterly averages of bank securities, loans, repos, and "other" interest-earning balances. Figure 7 decomposes aggregate interest-earning assets into these subcomponents over time, as a share of aggregate interest-earning assets from 2001-2022.[3] Loans typically represent the largest share of interest-earning assets followed by securities, averaging 58 percent and 21 percent over the sample period. From 2019 Q4-2022 Q1, loans' share of aggregate interest-earning assets decreased by 10 percentage points and was reallocated toward securities and other interest-earning assets, the latter being largely composed of balances due from depository institutions. While these portfolio changes occurred in aggregate, Figure 8 shows that there is a large amount of cross-sectional heterogeneity in the percentage of interest-earning assets represented by the loan and security portfolios across banks over time,

[3] In the FR Y 9-c reporting forms, the "other" category of interest-earning assets is not reported until 2001.

suggesting a wide variety of business models. By 2021 Q1, loans ranged from 31 to 87 percent of interest-earning assets between the top and bottom 5th percentiles, and securities ranged from 5 to 41 percent of interest-earning assets between the top and bottom 5th percentiles.

In the period following the pandemic, banks' loan balances declined while strong deposit growth was invested into investment securities and cash balances, which offer relatively lower rates of return.[4] Despite the reduction in loans as a share of interest-earning assets after the onset of the pandemic period, commercial and industrial loans did experience significant but temporary growth beginning in 2020 Q1 from increased drawdown of business lines of credit and the Paycheck Protection Program (PPP) lending that occurred as part of the federal response to the pandemic (Ennis and Jarque (2021)). The increase in C&I lending offered relatively lower rates of return, as the PPP loans all carried an interest rate of 1 percent and the commercial and industrial loan drawdowns would be priced during a period of falling interest rates.[5]

These summary statistics are suggestive of bank heterogeneity in the distribution of net interest margins that may correlate with skewness observed in the tails of the distribution, and this heterogeneity may be more pronounced during the post-pandemic period. Heterogeneity in sensitivities to interest rates and credit spreads speaks to variation in bank sensitivity to macroeconomic risks, and heterogeneity in the persistence of bank NIM similarly speaks to the degree to which past shocks to NIM influence future outcomes. Given (i) bank NIM represents profits related to macroeconomic interest rates and credit risk, (ii) the importance of persistence in explaining bank profits, and (iii) observable bank heterogeneity across the NIM distribution, I ask the questions: is there heterogeneity in sensitivity to interest rates, credit risk, and persistence across the conditional NIM distribution, and is this heterogeneity more pronounced in the post-pandemic period?

---

[4]As discussed in the April 2021 publication of the Federal Reserve Board's Supervision and Regulation Report. See https://www.federalreserve.gov/publications/2021-april-supervision-and-regulation-report-banking-system-conditions.htm.

[5]PPP loan details are defined by the SBA here: https://www.sba.gov/funding-programs/loans/covid-19-relief-options/paycheck-protection-program/first-draw-ppp-loan.

4 Empirical Framework

I follow Covas, Rump, and Zakrajšek (2014) to estimate a dynamic panel quantile regression model of bank NIM. The focus of this analysis is bank NIMs' distributional sensitivities to macroeconomic interest rates, credit risk, and own persistence. In this setting, interest rates and credit risk are treated as exogenous and the focus is on relationships between observed macroeconomic fluctuations and bank NIMs. In a linear setting, dynamic panel model is specified as

$$Y_{i,t} = \alpha_i + \sum_{s=1}^{k} \phi_s Y_{i,t-s} + \boldsymbol{X}'_{i,t-1}\boldsymbol{\beta} + \boldsymbol{Z}'_t\boldsymbol{\gamma} + \epsilon_{i,t}, \quad i = 1, ..., n, \; t = 1, ..., T, \; N = n \times T \tag{1}$$

where Y is NIM, or one of its components, realized for bank i at time t, α_i is a vector of individual intercepts meant to control for unobserved heterogeneity, X is a vector of balance sheet controls for time-varying heterogeneity across banks over time, and Z is a vector of macroeconomic interest rate and credit risk factors, including the level of short-term interest rates, the slope of the yield curve, and the credit risk spread.

Quantile regression provides a natural framework for examining conditional distributional sensitivities and allows for heterogeneous impacts of covariates across the conditional distribution. As one approach to measuring risk, this framework is motivated by its ability to capture nonlinearities in conditional outcomes, relevant during periods of severe macroeconomic and financial stress, and to recover the entire conditional distribution of the outcome rather than the conditional mean. The conditional quantile function of $Y_{i,t}$ in equation 1 is

$$Q_{Y_{i,t}}(\tau_k | \alpha_i, Y_{i,t-1}, ..., Y_{i,t-k}, \boldsymbol{X}'_{i,t-1}, \boldsymbol{Z}'_t) = \alpha_i + \sum_{s=1}^{k} \phi_s Y_{i,t-s} + \boldsymbol{x}'_{i,t-1}\boldsymbol{\beta} + \boldsymbol{Z}'_t\boldsymbol{\gamma} \tag{2}$$

where τ_k is a quantile in (0, 1).

Estimation of individual effects in quantile regression is complicated by the fact that the use of the expectations operator is no longer valid, leading to the incidental parameters problem that can bias the remaining parameter estimates of interest. Koenker (2004) proposes a penalized estimation approach to deal with the incidental parameters problem associated with estimating many fixed effects in a nonlinear framework. To reduce this bias, I place a LASSO shrinkage on the individual effects which improves the consistency of the remaining parameter estimates. LASSO estimation penalizes parameter estimates by the sum of their absolute magnitude, and results in some coefficients being shrunk exactly to zero resulting in a sparse solution (Tibshirani (1996)). I apply this to the incidental parameter problem present in my empirical specification, and shrinkage is achieved via a penalty term placed on the individual effects. Regression parameters are thus estimated as the solution to the problem

$$\underset{\beta\in\mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^{n}\sum_{t=1}^{T} w_\tau \rho_\tau(y_{it} - \alpha_i - \sum_{s=1}^{k}\phi_s Y_{i,t-s} - \boldsymbol{x}'_{i,t-1}\boldsymbol{\beta} - \boldsymbol{Z}'_t\boldsymbol{\gamma}) - \lambda\sum_i |\alpha_i| \tag{3}$$

where the weight w_τ controls the relative influence of the τth quantile on estimation of α_i, $\rho_\tau(u) = u\{\tau - I(u < 0)\}$ is the so-called "pinball loss" function used in quantile regression, and the final term is the LASSO penalty with shrinkage parameter λ. As $\lambda \to \infty$, the LASSO penalty shrinks individual intercepts toward zero.

4.1 Bayesian Panel Quantile Regression with Adaptive LASSO

Differently than Covas, Rump, and Zakrajšek (2014), I use a flexible Bayesian model specification that employs adaptive LASSO shrinkage on the individual effects. When the dimensions of the design matrix used to estimate (3) are large, the optimization exercise may be costly. LASSO shrinkage is well suited to the Bayesian paradigm, resulting from a specific prior placed on the parameters subject to the LASSO penalty.

In the frequentist tradition, the conditional quantile function is given by

$$Q_{Y_{i,t}}(\tau_k|\mathbf{X}'_{it}) = \mathbf{x}'_{it}\boldsymbol{\beta},$$

and regression coefficients are estimated as the solution to

$$\underset{\beta}{\operatorname{argmin}}\hat{\beta}_\tau = \sum_{i=1}^{n}\sum_{t=1}^{T}\rho_\tau(y_{it} - \mathbf{x}'_{it}\boldsymbol{\beta}) \tag{4}$$

The Bayesian approach to quantile regression emphasizes the link between minimizing

(4) and maximizing the likelihood of the asymmetric Laplace distribution (ALD) having location, scale, and skewness parameters μ, σ and τ, respectively, so that $y_{ij} \sim ALD(\mu, \sigma, \tau)$ (Yu and Moyeed (2001)). In this case, setting mean $\mu_{it} = \mathbf{x_{it}}'\boldsymbol{\beta}$ with $\mathbf{X} = \mathbf{x}_{11}, ..., \mathbf{x}_{nT}$ and $y = y_{11}, ..., y_{nT}$ and assuming $y_{it} \sim ALD(\mu_{it}, \sigma, \tau)$, the likelihood across $n \times T$ independent observations is

$$L(\boldsymbol{\beta}, \sigma|\boldsymbol{y}, \boldsymbol{X}) = \frac{\tau(1-\tau)^N}{\sigma^N} exp\{-\sum_{i=1}^{n}\sum_{t=1}^{T} \rho_\tau \frac{(y_{it} - \boldsymbol{x}'_{it}\boldsymbol{\beta})}{\sigma}\} \tag{5}$$

Yu and Moyeed (2001) show that maximization of likelihood (5) is equivalent to the minimization of (4).

I follow Aghamohammadi and Mohammadi (2017), who propose a Bayesian approach to estimate (3) via a tractable Gibbs sampling algorithm for posterior inference in a random intercept hierarchical framework. The Bayesian literature typically implements a Laplace prior for LASSO regularization of parameter estimates (for example, see Park and Casella (2008)). Aghamohammadi and Mohammadi (2017) show that, under certain assumptions, if we place a Laplace prior on random intercepts $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_n)$, resulting in $\pi(\boldsymbol{\alpha}|\lambda, \sigma) = (\frac{\lambda}{2\sqrt{\sigma}})^n exp\{-\frac{\lambda}{\sqrt{\sigma}} \sum_{i=1}^{n} |\alpha_i|\}$, and further assume $\nu = \frac{\lambda}{\sqrt{\sigma}}$, then the posterior distribution of $\boldsymbol{\alpha}$ can written as

$$\pi(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \sigma, \lambda) \propto \sigma^{-N}\nu^n exp\{-\sum_{i=1}^{n}\sum_{t=1}^{T} \rho_\tau \frac{(y_{it} - \alpha_i - \sum_{s=1}^{k} \phi_s Y_{i\ t-s} - \boldsymbol{x}'_{i\ t-1}\boldsymbol{\beta} - \boldsymbol{Z}'_t\boldsymbol{\gamma})}{\sigma}\}$$
$$\times\ exp\{-\nu \sum_{i=1}^{n} |\alpha_i|\}. \tag{6}$$

Considering σ and λ as nuisance parameters equates maximization of (6) to minimization of (3). To obtain an efficient Gibbs sampler, Park and Casella (2008) specify the Bayesian LASSO by writing the Laplace distribution as a scale mixture of normals,

$$\pi(\boldsymbol{\alpha}|\lambda, \sigma) = \prod_{i=1}^{n} \int_0^\infty \frac{1}{\sqrt{2\pi s_i}} exp\left\{-\frac{\alpha_i^2}{2s_i}\right\} \times \frac{\nu^2}{2} exp\left\{-\frac{\nu^2 s_i}{2}\right\} ds_i,$$

I follow Aghamohammadi and Mohammadi (2017) in employing exponential priors on ν^2 of the form $\pi(\nu^2|\phi) = \phi\ exp\{-\phi\nu^2\}$, where $\phi > 0$ is a hyperparameter. Similarly motivated by the use of a Gibbs sampling algorithm, Kozumi and Kobayashi (2011) show that the ALD likelihood can also be treated as a scale mixture of normal distributions, so that assuming

$$y_{it}|e_{it}, \sigma \sim N(\alpha_i + \sum_{s=1}^{k} \phi_s y_{i\ t-s} + \mathbf{x}'_{i\ t-1}\boldsymbol{\beta} + \mathbf{z}'_t\boldsymbol{\gamma} + (1-2\tau)e_{ij},\ 2\sigma e_{ij})$$

results in $y_{ij} \sim ALD(\alpha_i + \sum_{s=1}^{k} \phi_s y_{i\ t-s} + \mathbf{x}'_{i\ t-1}\boldsymbol{\beta} + \mathbf{z}'_t\boldsymbol{\gamma},\ \sigma,\ \tau)$ when e_{ij} has an exponential distribution with mean $\frac{\sigma}{\tau(1-\tau)}$. The remaining parameters are assumed to be mutually independent with the following priors:

$$\beta \sim N_k(\mathbf{b}_0, \mathbf{B}_0), \quad \gamma \sim N_j(\mathbf{g}_0, \mathbf{G}_0), \quad \sigma \sim IG(c_0, d_0), \quad \pi(\phi) \propto \frac{1}{\phi}$$

Additional flexibility is achieved by incorporating the adaptive LASSO, used to deal with bias in parameter estimates and inconsistent variable selection in stemming from the fixed λ penalty term placed on all the individual effects in the traditional LASSO estimation (Zou (2006)). This results in a modification to the the Laplace prior distribution for α_i, where

$$\pi(\alpha_i|\lambda, \sigma) = (\frac{\lambda_i}{2\sqrt{\sigma}})exp\{-\frac{\lambda_i}{\sqrt{\sigma}}|\alpha_i|\},$$

and assuming $\nu_i = \frac{\lambda_i}{\sqrt{\sigma}},\ i = 1, ..., n$. This flexible framework allows the LASSO penalty term λ_i to vary across the individual effects.

Aghamohammadi and Mohammadi (2017) propose the following hierarchical model,

$$\begin{aligned} y_{it}|e_{it}, \beta, \sigma, \alpha_i \sim N(\alpha_i + \sum_{s=1}^{k} \phi_s y_{i\ t-s} + \mathbf{x}'_{i\ t-1}\boldsymbol{\beta} + \mathbf{z}'_t\boldsymbol{\gamma} + (1-2\tau)e_{it}, 2\sigma e_{it}), \\ e_{it}|\sigma \sim Exp(\frac{\tau(1-\tau)}{\sigma}), \\ \boldsymbol{\beta}|\mathbf{b}_0, \mathbf{B}_0 \sim N_k(\mathbf{b}_0, \mathbf{B}_0), \\ \boldsymbol{\gamma}|\mathbf{g}_0, \mathbf{G}_0 \sim N_j(\mathbf{g}_0, \mathbf{G}_0), \\ \alpha_i|s_i \sim N(\alpha_i, s_i), \quad s_i|v_i^2 \sim Exp(\frac{\nu_i^2}{2}), \nu^2|\phi \sim Exp(\phi), \pi(\phi) \propto \frac{1}{\phi}, \\ \sigma|c_0, d_0 \sim IG(c_o, d_0), \end{aligned} \tag{7}$$

where $N(\cdot)$ represents a Normal distribution, $Exp(\cdot)$ represents an exponential distribution, and $IG(\cdot)$ represents the Inverse Gamma distribution. This results in a tractable Gibbs sampling algorithm that relies on closed-form full conditional distributions.

Henceforth, I refer to the model as bayesian panel quantile regression with adaptive LASSO (BPQR-AL). I use the Gibbs sampling algorithm proposed by Aghamohammadi and Mohammadi (2017) to estimate the hierarchical framework and analyze the effects of interest rates, credit risk, and persistence on bank net interest margins while controlling for time-varying and time-invariant bank heterogeneity and reducing bias from the incidental parameters problem. The use of adaptive LASSO offers additional flexibility in cases where unobserved bank heterogeneity for individual banks offers relatively more information across the conditional distribution than other banks, thus allowing the random effects to explain more of the bank-level variation in such cases. In each set of results, I take 30,000 draws from the Gibbs sampler and discard the first 10,000 as the burn-in samples. I use the following weak priors on the parameters,

$$\boldsymbol{\beta} \sim N_k(0, 100I), \quad \boldsymbol{\gamma} \sim N_j(0, 100I), \quad \sigma \sim IG(0.01, 0.01).$$

5 Data

To examine sensitivities of bank NIM to interest rates, I collect quarterly data on bank holding companies balance sheet regulatory filings from 1998-2022 contained in form FR Y-9c that have been adjusted for mergers. The merger-adjustment is common in the literature used to forecast NIM, and necessary because bank mergers may change the way that income is reported in regulatory filings. The merger-adjusted data deal with this by accounting for mergers in the acquiring institution occurring during the calendar year prior to the actual merger taking place. These adjustments deal with reporting issues that may result in income data ultimately not being reported for calendar years in which banks become involved in a merger, which could bias estimates of income correlations with other covariates.[6]

[6]The bank filing data are merger-adjusted in accordance with methods described in the appendix of William B. English and William R. Nelson (1998), “Profits and Balance Sheet Developments at U.S. Commercial

Using these data, I construct an unbalanced panel of top-holder US bank holding companies (banks) from 1998-2022 having an average of at least $500 million in real total assets over the entirety of the the sample period. I choose this asset threshold for two reasons. First, $500 million was the relevant asset threshold for banks that would become subject to the US implementation of the Basel III regulatory standards in the post-GFC period.[7] Second, this asset threshold partly deals with population changes due to bank asset growth and consolidation, as well as secular trends associated with inflation.

While this threshold excludes some smaller banks early in the sample period, additional sample attrition occurs as a result of changes in the size requirements of FR Y-9c filers that occurs later in the sample period. Prior to March 2006, banks with at least $150 million in total assets were required to file the FR Y-9c form. This changed three times over the sample period: first from $150 million to $500 million in March 2006, next from $500 million to $1 billion in March 2015, then from $1 billion to $3 billion in September 2018. These changes were a result of inflation, consolidation, growth of asset balances, and reduction of bank burden.[8] Despite these changing requirements, some banks below the asset thresholds continued to file the FR Y-9c form after the threshold changes occurred. During the 2015 threshold change, the sample size decreased by 8 banks between December 2014 and March 2015. During the 2018 threshold change, the sample size decreased by 47 banks between June 2018 and September 2018. I treat these population changes as exogenous and do not make further sample selection adjustments since they do not impact the primary point of comparison in parameter estimates between the pre-2020 and post-2020 periods. I do, however, examine results based on bank size subsamples that would exclude impacts of the changing reporting thresholds for the relatively larger banks.

In addition to the issue of incidental parameters in the dynamic panel quantile regression model, there is also a problem of bias in parameter estimates that arise from inclusion of lagged dependent variables in linear and quantile regression frameworks using "big N, small T" panel data (Nickell (1981); Galvao (2011)). I follow Covas, Rump, and Zakrajšek (2014)

Banks in 1997," Federal Reserve Bulletin, vol. 84 (June), p. 408.

[7]The US implementation of the Basel III regulatory standards applied to banks with at least $500 million in total assets, with additional enhanced prudential regulations for the largest institutions. See https://www.govinfo.gov/content/pkg/FR-2013-10-11/pdf/2013-21653.pdf

[8]See https://www.federalreserve.gov/apps/reportingforms/Report/Index/FR_Y-9C.

in estimating the model on a relatively long time series for the panel of banks in the sample such that each bank must be observable for at least 10 years to be included in the sample; therefore the effect of the bias is insignificant, since the initial conditions will have little impact on parameter estimates.[9]

The remaining sample selection and data cleaning criteria are as follows. I remove any observations with values equal to zero for total assets or interest-earning assets. I also remove any observations in the data that have missing values for variables included in the empirical model. The described data cleaning results in an unbalanced panel of 313 unique bank holding companies and 24,326 bank-quarter observations spanning 1998-2022. This sample, on average, represents 85 percent of total assets across all bank holding companies in the merger-adjusted data.

I focus on the yield curve factors (level and slope) and a credit spread as the relevant macroeconomic variables. I use the market yield on Treasury securities at 3-month constant maturity to measure the level of the yield curve, and the difference between the market yields on 10-year and 3-month Treasury securities at constant maturity to measure the slope of the yield curve. These data are collected from the H.15 interest rate data published by the Federal Reserve Board of Governors. To construct a credit spread, I use the difference between the ICE 7-10 year US corporate index effective field and the 10-year Treasury constant maturity rates.

Aside from macroeconomic interest rates, credit risk, and persistence, the theoretical and empirical banking literature suggests additional bank characteristics that can affect NIM; I use these characteristics as control variables in the quantile regression analysis. In the micro literature, NIM of a profit-maximizing bank is a function of, among other things, the bank's risk aversion, business model, market power, credit risk and interest rate risk premiums (for example, Ho and Saunders (1981), McShane and Sharpe (1985), Allen (1988), Angbazo (1997), Hanweck and Ryu (2005), Saunders and Schumacher (2000)). Proxies for these variables vary in the literature, but include bank capital ratios as proxies for risk aversion, measures of bank size or market share for market power, bank balance sheet concentrations for business model, and balance sheet credit losses for bank-specific credit risk. Other empirical literature

[9] In fact, imposing this filter on my sample results in the average bank being observed for roughly 20 years.

cites maturity mismatch (Flannery (1981), English, Van Den Heuvel, and Zakrajšek (2018)), reliance on relatively cheap "core" deposits, and the use of interest rate derivatives (English, Van Den Heuvel, and Zakrajšek (2018), Drechsler, Savov, and Schnabl (2021)) as bank-specific determinants of NIM.

The empirical proxies that I include to capture these characteristics for each bank include Tier 1 leverage ratio, log of real total assets, loans as a share of interest-earning assets, securities as a share of interest-earning assets, gross notional interest rate derivatives share of assets, net charge-offs as a share of total loans, core deposits as a share of liabilities, and the maturity gap proposed by English, Van Den Heuvel, and Zakrajšek (2018).[10] Table 1 presents summary statistics for all of the variables used in the empirical analysis.

6 Estimation Results

6.1 Full Sample Results

As a first point of comparison, I estimate fixed-effects OLS (FE_OLS) parameters of model 1 in addition to the BPQR-AL parameter estimates. Table 2 shows the FE-OLS and BPQR-AL results, estimated across all banks in the sample over the period from 1998 Q1-2022 Q4. All the estimates shown include individual bank effects and the set of balance sheet controls described previously.

Column 1 shows the FE-OLS model parameter estimates and 95 percent confidence intervals. The sum of the autoregressive terms in the model is roughly 0.81 and statistically significant, suggesting a high level of persistence in bank NIM. The results imply that there is a positive and statistically significant correlation between the level and slope of the yield curve factors and bank NIM, consistent with findings in the previous literature. These estimates suggest that 100 basis point increases in the 3-month Treasury bill and Treasury spread

[10] See the appendix of English, Van Den Heuvel, and Zakrajšek (2018) that details construction of the maturity gap, https://ars.els-cdn.com/content/image/1-s2.0-S0304393218302101-mmc1.pdf. For a sample of bank holding companies, this requires aggregating Call Report data up to the top-holder bank holding company level. I follow Drechsler, Savov, and Schnabl (2021) in assuming that (i) balance sheet items without maturity information and extremely short-maturity items such as transaction and savings deposits, cash, fed funds bought or sold, and repo have a maturity of zero, and (ii) subordinated debt has a repricing maturity of 5 years.

rates result in bank NIMs increasing by roughly 7 and 9 basis points on average, respectively. There is a negative correlation of the credit spread with NIM, although the effect is close to zero in size and is not statistically significant.

Columns 2-6 show the posterior means and 95 percent credible intervals of the BPQR-AL model parameters at the 5th, 25th, 50th, 75th and 95th conditional quantiles.[11] The sum of autoregressive terms in the model show that, while persistence is high across all quantiles, it is relatively lower at the 5th and 95th quantiles and higher near the median so that shocks to NIM that move banks toward the tails of the distribution tend to be less persistent. Sensitivity to the yield curve factors is both relatively highest at the 5th and 95th quantiles, and relatively lower near the median. A 100 basis point increase in the 3-month Treasury Bill is associated with roughly 5, 3, and 8 basis point increases in NIM at the 5th, 50th and 95th conditional quantiles, respectively. A 100 basis point increase in the Treasury spread is associated with roughly 7, 5, and 9 basis point increases in NIM at the 5th, 50th and 95th conditional quantiles, respectively. Sensitivity to the credit risk spread moves from negative to positive across the conditional quantiles, so that a 100 basis point increase in the credit risk spread is associated with a 4 basis point decrease in NIM at the 5th conditional quantile and a 4 basis point increase in NIM at the 95th conditional quantile. These results suggest that interest rate and credit risk factors have relatively more influence at the conditional tails of the distribution, while the conditional median of the distribution is more strongly influenced by past realizations.

Figure 9 provides a graphical comparison of the BPQR-AL and FE-OLS estimates for the interest rate and autoregressive variables. The figure shows that the FE-OLS estimates generally appear to do a poor job of fitting the estimated median of the conditional distribution, often fitting closer to the tails. As a result, the FE-OLS estimates may overstate the impact of the interest rate factors across the central tendency of the conditional distribution, while impact of the sum of autoregressive terms (and, to a lesser extent, the tail effects of the credit risk factor) is understated. This would cause conditional mean forecasts to be more responsive to contemporaneous interest rate shocks relative to conditional median forecasts.

[11] See the Appendix for trace and ACF plots used to monitor the convergence of the Gibbs sampling algorithm used for the BPQR-AL model.

The BPQR-AL model, however, suggests that the tails are relatively more responsive to interest rate and credit risk factors, while the median of the distribution is relatively more responsive to own lags and less so to macroeconomic variables.

Seemingly small, the BPQR-AL parameter estimates in the tails of the conditional distribution can be economically meaningful when interest rate movements are large, as they were during 2020-2022. For example, from 2021 Q4 to 2022 Q4, the quarterly average 3-month Treasury bill rate increased by roughly 400 basis points as monetary policy rates reacted to rising inflation. Over this period, the aggregate impact of the 3-month Treasury rate increase would sum to an additional 5.6 basis point increase in NIM at the 5th conditional quantile relative to the median, and a 20 basis point increase at the 95th conditional quantile relative to the median. In 2021 Q4, the unconditional lower 5th percentile of the NIM distribution was 1.17 percent of interest-earning assets, and for these banks an additional 5.6 basis points is approximately 5 percent of total NIM. The magnitude of this difference is similar for the top 95th percentile.

To demonstrate the impact of the adaptive LASSO in the BPQR-AL, Figure 10 shows the posterior mean estimates of the individual bank random effects α_i relative to the posterior mean estimate of the shrinkage parameter ν_i at each estimated quantile. The pattern that is exhibited in each plot resembles a Laplace distribution, reflecting the Bayesian implementation of the LASSO shrinkage placed on the individual effects. Because the amount of shrinkage is allowed to vary across each bank, the prior assumptions placed on the individual effects in conjunction with the data cause estimated effects that are relatively small (large) to have a higher (lower) amount of shrinkage placed on them, so that larger values of ν_i are associated with values of α_i that are closer to zero and thus could be discarded by the econometrician.[12] As the model allows for individual effects that vary across the conditional distribution, these plots suggest that some individual effects become relatively large in absolute magnitude in the tails so that unobserved heterogeneity can be important for certain banks that experience tail outcomes.

[12]There exist rules of thumb for choosing which effects to discard in a Bayesian LASSO setting in order to achieve a more parsimonious specification. One approach is to use credible intervals to guide variable selection, see Park and Casella (2008).

6.2 Time Subsample Results

The literature on bank NIMs documents increased and heterogeneous sensitivity to interest rate risk in the post-2020 period as a result of increases in the share of long-term securities held by banks, increased non-deposit funding costs during the monetary policy tightening period beginning in 2022, and differences in bank size (Abdymomunov, Gerlach, and Sakurai (2023), Laliberte and Sengupta (2024), Sengupta and Xue (2022)). Based on these documented results, it may be that the BPQR-AL model would show increased sensitivities to interest rates in a sample including the post-2020 period relative to a sample restricted to earlier periods, and that these increased sensitivities may vary by bank size across the conditional distribution.

To examine changes in sensitivities in the pre-2020 and post-2020 periods, I estimate the BPQR-AL model on banks in the sample spanning 1998-2019 for comparison to the estimates previously shown for the full 1998-2022 sample. Figure 11 plots the BPQR-AL posterior estimates for the autoregressive, interest rate, and credit risk variables by the time period included in the estimation sample. Most notably, the results show that inclusion of the 2020-2022 period in the estimation sample results in increased sensitivity to both of the interest rate factors across the conditional quantiles, and the increase is relatively more pronounced at both tails of the conditional distribution for the 3-month Treasury bill and the lower tails for the Treasury spread. The top-right plot shows that estimated 3-month Treasury sensitivities at the lower 5th quantile and median were roughly equivalent in the 1998-2019 sample. In the 1998-2022 sample, sensitivity at the lower 5th quantile increased by 2 basis points, while the median sensitivity increased by less than 1 basis point.

The remaining coefficients show relatively more stability across the conditional quantiles in the subsample time periods. The top-left plot shows that the posterior mean estimates for persistence decreased slightly at the upper tail of the conditional distribution; however, there is large overlap in credible intervals of posterior mean persistence estimates in the time subsamples. The bottom-right plot shows increased sensitivity to the credit spread at the upper conditional quantiles, while the estimated sensitivities remain similar across the median and lower quantiles. Given the smaller amount of variation in the credit spread relative to the interest rate factors, particularly the 3-month Treasury bill rate, this increased sensitivity may have a smaller impact on the overall quantile forecasts of NIM.

6.3 Income and Expense Decomposition

The time variation of interest rate sensitivities may in part reflect changes to income and expenses resulting from balance sheet dynamics that occurred during the post-2020 period, as described in Section 3. To further examine the source of time variation in sensitivities to interest rate factors, I decompose the numerator of net interest margin into the constituents of interest income and interest expense. Interest income is decomposed into income earned on loans, securities, and all other sources of interest income. Interest expense is decomposed into expenses paid on demand deposits, time deposits, trading expenses, and all other sources of interest expense. For the components of interest income and interest expense, I construct a ratio using interest-earning assets as the denominator so that differences in the composition of bank's balance sheets do not impact the results. I then estimate the BPQR-AL model over the two time period subsamples, replacing the dependent variable with each component's share of interest-earning assets, maintaining an AR(4) specification along with the same set of lagged balance sheet characteristics and contemporaneous interest rate and credit risk factors.

Because the quantile regression estimates are nonlinear in nature, a net effect of the time variation observed in the decomposition results is not straightforward to infer. Despite this complexity, it is reasonable to assume that the time variation of sensitivity to interest rates observed in the lower tail of the interest income components and the upper tail of the interest expense components may correlate with the time variation observed in the lower tail sensitivities of bank NIM.

Figure 12 shows interest rate factor estimates from the interest income decomposition. Dependent variables as a share of interest-earning assets are in the columns, and the associated interest rate factor estimates are in the rows. Within the plots, each line shows the quantile-specific estimates for the component of interest income in the two subsample time periods. For readability, I do not show the credit risk or bank balance sheet parameter estimates, although these were included in the estimations. The figure shows that interest income on loans as a share of interest-earning assets explains the largest increase in sensitivity to both interest rate factors relative to securities and other interest income. Sensitivities are increased for loan income across the conditional quantiles, and increase by relatively more at the lowest conditional quantile estimates. In terms of magnitudes, the tail sensitivities of the loan portfolio are larger for the short-term rate when compared to the Treasury spread.

Figure 13 shows the decomposition for interest expense components, which suggests that demand deposits explain most of the increase in interest expense sensitivity. This increase is particularly pronounced for the 3-month Treasury rate at the upper tails of the conditional distribution. In terms of magnitudes, the size of demand deposits' sensitivities to the 3-month Treasury rate is notably smaller than those observed for loan income. Sensitivities for time deposits show the opposite pattern, decreasing in the subsample that includes the post-2020 period, although the size of the decrease is relatively small across all of the conditional quantiles. In all cases, quantile sensitivities to the Treasury spread are relatively low, as expected for relatively short-term bank liabilities.

6.4 Bank Size Results

Another dimension by which interest rate sensitivity varies over time is bank size. Studies have documented NIM differentials in terms of bank size (Abdymomunov, Gerlach, and Sakurai (2023), Sengupta and Xue (2022)). Figure 14 shows that the largest (smaller) banks tend to have relatively lower (higher) NIM, sometimes by as much as roughly 1 percentage point. One explanation for this is that larger commercial banks tend to have relatively higher proportions of non interest income, which could offset lower net interest incomes (Haubrich and Young (2019)). I estimate the BPQR-AL model on persistence, interest rate, and credit risk factors, and the balance sheet controls described in Section 4 across the conditional distribution of NIM for banks of different size groupings. I segment the sample of banks into three size groups based on average real total assets over the sample period to remove any noise due to banks moving between thresholds during the sample period. The size groups include banks less than \$10 billion, \$10 billion to \$100 billion, and greater than \$100 billion in real total assets. These categories are related to common US definitions of community, regional, and large banks.

Figure 15 shows the BPQR-AL posterior estimates for the autoregressive interest rate and credit risk variables by bank size grouping. The posterior mean estimates of the sum of autoregressive terms for the larger banks are relatively higher at the lower tail of the distribution, and relatively lower at the higher tail of the distribution, but the 95 percent credible interval around estimates for the largest banks is very wide in this case, encompassing the credible intervals of the smaller size groups. Larger banks tend to be slightly more sensitive to the level of the yield curve in the conditional tails of the distribution, and slightly more sensitive to the Treasury spread at the highest tail of the conditional distribution. Median interest rate sensitivity estimates across the size groups are roughly similar, but differences do appear in the conditional tails. A 100 basis point increase in both the 3-month Treasury bill rate and the Treasury spread is associated with an additional 4basis point (5 basis point) increase in NIM for the largest banks relative to the smallest banks at the 5th (95th) conditional quantile. Finally, the largest banks appear to be relatively more sensitive to the credit risk spread at the median and higher tail of the conditional distribution, and the relationship is positive, which suggests that these banks may do a better job of pricing credit risk. A 100 basis point increase in the credit risk spread, slightly larger than a one standard deviation movement, is associated with a roughly 6 basis point increase in NIM for the largest banks relative to the smallest banks at the 95th conditional quantile.

Given the time variation in interest rate sensitivities shown previously, it is useful to examine how the time variation changes across the size distribution. Figure 16 shows the BPQR-AL posterior estimates for the interest rate factors by time period subsample and bank size groups. Two interesting findings emerge from these results. First, regardless of time period, one primary difference in the sensitivities of the interest rate factor lies in the upper tails of the conditional distribution where the larger banks generally exhibit a higher level of sensitivity relative to the smallest banks. Second, the increase in lower tail sensitivities in the 1998-2022 subsample is relatively largest for the largest banks. For banks greater than \$100 billion in total assets, sensitivity to a 100basis point increase in the 3-month Treasury bill increased from 0.04 to 0.09 basis points between the subsamples ending in 2019 and 2022, respectively, and sensitivity to a 100 basis point increase in the Treasury spread increased from 0.03 to 0.11 basis points. These changes are roughly 4-5 times greater than the sensitivity changes for the smallest banks. This suggests that the largest banks that drove the overall increase in tail sensitivity across the conditional distribution in the post-2020 period.

7 Density Forecasting

In this section, I compare density forecasts across competing models to test if the features of the BPQR-AL model meaningfully improve the model predictions relative to alternative model approaches. The variation in tail sensitivities over time and across banks suggests that a bank NIM model specification that allows for heterogeneous parameter estimates should be better suited to capturing the tails of the conditional distribution relative to the conditional mean estimation. Additionally, the quantile-varying individual effects and adaptive LASSO shrinkage may improve model predictions relative to the frequentist analog of the model used by Covas, Rump, and Zakrajšek (2014) that relies on a simpler LASSO framework applied to individual effects that are constant across quantiles.

I evaluate the density forecast performance of the BPQR-AL model relative to fixed-effect OLS (FE-OLS) and pooled quantile regression (QR) benchmark specifications of equation (1). Similar to Covas, Rump, and Zakrajšek (2014), I also estimate a frequentist fixed-effects quantile autoregressive model (FE-QAR) specification based on equation (3), which imposes a single LASSO shrinkage parameter across quantile-invariant individual effects.[13] In each case, I use the same set of model covariates reported in the previous estimation results. If the tails of the distribution matter, if individual effects capture unobserved bank heterogeneity, and if adaptive LASSO shrinkage on the individual effects improves upon model forecasts then the BPQR-AL model should outperform FE-OLS, QR, and FE-QAR methods.

7.1 Density Forecast Construction

I evaluate in-sample and pseudo out-of-sample forecasts for competing model specifications. First, I estimate each model over the period 1998-2019, which I treat as the in-sample period.

[13]In the FE-QAR model estimation, I follow Covas, Rump, and Zakrajšek (2014) in setting the single shrinkage parameter λ equal to 1.

Next, I produce pseudo out-of-sample one-quarter-ahead quantile forecasts for each bank by updating the lagged bank characteristics and contemporaneous interest rate and credit risk factors with data from the subsequent quarter, where the first out-of-sample quantile forecast for each bank is computed from the existing set of in-sample model parameters.[14] Then the model parameters are updated by estimating the updated set of dependent variable realizations on the updated covariate data, after which the data are again updated, and the subsequent out-of-sample quantile forecasts for each bank are produced. This process repeats until the final forecast is produced in 2022 Q4.

Once the quantile forecasts are obtained, there are several possible approaches to constructing density forecasts in the academic literature. For the BPQR-AL, QR, and FE-QAR models, I follow the approach of Adrian, Boyarchenko, and Giannone (2019) and Giglio et al. (2021) in fitting the estimated conditional quantile forecasts produced to the skewed t-distribution proposed by Azzalini and Capitanio (2003). This approach is attractive because it allows for an estimate of the full conditional density using a parsimonious set of estimated conditional quantiles. The form of the skewed t-distribution is,

$$f_y(\mu, \sigma, \alpha, \nu) = \frac{2}{\sigma} t(\frac{y-\mu}{\sigma}; \nu)\ T\left(\alpha \frac{y-\mu}{\sigma} \sqrt{\frac{\nu+1}{\nu + (\frac{y-\mu}{\sigma})^2}}\right), \tag{8}$$

where $t(\cdot)$ and $T(\cdot)$ represent the probability density and cumulative distribution functions of the Student t-distribution, respectively. The skewed t-distribution is an extension of the Student t-distribution, where the parameter α regulates the skewness. The remaining distributional parameters μ, σ and ν represent the location, scale, and fatness, respectively.

I estimate one-quarter-ahead density forecasts for each bank in the sample over the in-sample and out-of-sample periods 1998-2019 and 2020-2022, respectively. For every bank in every time period in the sample, the four parameters μ, σ, α, and ν are chosen to minimize the squared distance between the estimated 5th, 25th, 75th and 95th conditional quantiles produced by the BPQR-AL model and the same quantiles of the skewed t-distribution. As noted by Adrian, Boyarchenko, and Giannone (2019), this is equivalent to an exactly identified

[14] While the model covariates include contemporaneous time (t) macroeconomic variables, this forecasting framework treats those as known values at time ($t-1$). Given that the bank data are quarterly and not made available until multiple weeks after each quarter-end, this forecasting framework is practical in design.

nonlinear cross-sectional regression of the predicted quantiles on the quantiles of the skewed t-distribution.

The FE-OLS model imposes normal distribution assumptions on the conditional density forecasts given an estimated mean and variance. Using the FE-OLS model, the one-quarter-ahead prediction for each bank is set as the mean of the forecast distribution at each point in time. One standard assumption of OLS with homoscedastic errors is that unobserved variance is assumed to be constant across observations, and is estimated from in-sample residuals resulting from each regression. The density forecasts for each bank during the in-sample period are produced using the estimated mean and variance from the in-sample OLS results. In the pseudo out-of-sample period, a density forecast for each bank is produced using the estimated mean and variance that are iteratively updated at each new estimation as the covariate data are updated.

7.2 Density Forecast Evaluation

To evaluate the performance of the density forecasts across the two models, I use the continuous ranked probability score (Gneiting and Raftery (2007)). A continuous ranked probability score (CRPS) that is closer to zero indicates a more accurate forecast density. The CRPS is computed by first computing the quantile score (QS), based on the pinball loss function that is fundamental to quantile regression. Assuming τ_p is a set of estimated quantiles p in $\{1, 2, ..., P\}$, given a quantile forecast f at time t for bank i, the QS is defined as

$$QS_{it}(\tau_p) = \begin{cases} (1-\tau_p)(f_{it}(\tau_p) - y_{it}) & \text{if } y_{it} < f_{it}(\tau_p) \\ \tau(y_{it} - f_{it}(\tau_p)) & \text{if } y_{it} \geq f_{it}(\tau_p) \end{cases} \tag{9}$$

Then, for a set of evenly spaced quantile estimates Q, the CRPS is computed as twice the average of the quantile scores,

$$CRPS_{it} = \frac{2}{P} \sum_{p=1}^{P} QS_{it}(\tau_p) \tag{10}$$

Gneiting and Ranjan (2011) propose a CRPS that places additional weight on particular regions of the density forecast to evaluate density forecasts with a focus on the tails or central region of the density forecast. They define the quantile-weighted continuous ranked probability score (qwCRPS) as

$$qwCRPS_{it} = \frac{2}{P}\sum_{p=1}^{P}\gamma(\tau_p)QS_{it}(\tau_p), \tag{11}$$

where the weighting function $\gamma(\tau_q)$ is defined to emphasize particular areas of the density forecast. To place more weight on the left tail of the density forecast, $\gamma(\tau_p)$ is set equal to $(1-\tau_p)^2$, and to place more weight on the right tail of the density forecast, $\gamma(\tau_p)$ is set equal to $(\tau_p)^2$. Finally, placing more weight on the center of the density forecast, $\gamma(\tau_p)$ is set equal to $\tau_p(1-\tau_p)$.

I extract 199 evenly spaced quantiles from the FE-OLS, QR, and BPQR-AL density forecasts of $\{0.005, 0.01, 0.015, ..., 0.995\}$ to compute the CRPS. To compare across models, I examine ratios of the average in-sample and out-of-sample scores across all banks and time periods for the alternative modeling frameworks using the density forecasts and actual outcomes for each individual bank. Ratios are computed as the average BPQR-AL score divided by the average score of a competing model, so that a number less (greater) than 1 indicates that the BPQR-AL model has a lower (higher) average score relative to the competing model in the given period. Assuming uncorrelated error terms across banks and time, one can use the Diebold-Mariano (DM) t-test of average loss equality across forecasts to statistically evaluate the differences in the CRPS metrics across model specifications for each bank in each time period (Gneiting and Katzfuss (2014); Granger and Huang (1997)). I use a one-sided Diebold-Mariano test as a formal test of relative model accuracy, where the alternative hypothesis is that the BPQR-AL model is more accurate than a competing model. In all DM test results, the model forecasts are treated as primitives, and heteroskedasticity and autocorrelation (HAC) robust standard errors are used for inference.[15]

[15] Despite the relatively short number of periods in the out-of-sample period, the panel data allow for greater statistical power in conducting this test.

7.3 Alternative Model Comparison Results

Table 3 reports ratios of the average score metrics, which include the average overall CRPS, and the average left, right, and center-weighted CRPS denoted as CRPS-L, CRPS-R, and CRPS-C, respectively. The table includes results of a one-sided DM test where the alternative hypothesis is that the BPQR-AL model is statistically more accurate than the competing model. During the in-sample period, the BPQR-AL has statistically significant lower scores across all metrics when compared to the FE-OLS model, and all but the CRPS-L metric for the FE-QAR model. The QR model appears to perform slightly better than the BPQR-AL model during the in-sample period; however, the DM test does not suggest any statistical improvements relative to the BPQR-AL model. In the out-of-sample period, the BPQR-AL model outperforms all competing model specifications across all score metrics, and in each case the DM test suggests that the BPQR-AL forecasts are statistically more accurate than those of the competing models. These results suggest that the BPQR-AL model provides a statistically more accurate density forecast during the out-of-sample period overall, and also in the tails and center region of the density forecasts for the individual banks. This provides additional evidence that estimation of the conditional distribution and individual effects using adaptive LASSO shrinkage improve upon density forecasts of bank NIM.

Figure 17 plots the CRPS across each model during each time period to examine where the accuracy gains occur for the BPQR-AL model relative to the alternative models. The BPQR-AL model outperforms the FE-OLS model in most time periods. The BPQR-AL and QR models perform similarly except in periods where the CRPS increases for all models, in 2020 Q2 and in all quarters during 2022, all periods of relatively large movements of the quarterly average 3-month Treasury rates included in the model. This suggests that the additional forecast uncertainty provided by the individual effects in the BPQR-AL model improves the forecast accuracy.

The individual bank density forecasts can be combined at each time period such that an aggregate density forecast is produced, which allows for a useful visual comparison of differences across the alternative model specifications. The aggregated forecasts may also be useful in a macro prudential setting where regulatory concerns lie around aggregate risks rather than bank-specific risks. Several methods exist for combining density forecasts. I follow Giglio et al. (2021) in applying Vincentization to the individual bank quantile forecasts to arrive at a distributional forecast of aggregate outcomes. Vincentization is a method of combining distributions by computing a weighted average of each distribution at the quantiles of each distribution. Specifically, assume $f_{it}(y_t)$ is the NIM density forecast of a given bank i at time t, and $F_{it}(y_t)$ is the cumulative distribution function for each bank i, $i = 1, ..., n$. Given a non-negative set of weights ω_i, where $\sum_{i=1}^{n} \omega_i = 1$, the combination of the conditional quantiles estimated is given by

$$\sum_{i=1}^{n} \omega_i F_{it}^{-1}(\tau), \quad 0 < \tau \leq 1$$

where $F^{-1}(\tau)$ represents the quantile functions of the individual banks in the sample (see Busetti (2017) for additional details on Vincentization).

I use Vincentization to combine the individual bank density forecasts into an aggregate forecast for comparison to the average NIM outcome across all banks in the sample at each time period. For simplicity, I set the ω_i such that the conditional quantiles are equally weighted. Figure 18 shows the aggregated density forecasts along with the average realization of NIM across all banks in the sample at each time period. As before, the in-sample period is 1998-2019, and the pseudo out-of-sample forecasts begin in 2020 Q1 and continue through 2022 Q4. Relative to the QR and BPQR-AL models, the FE-OLS model produces one-quarter-ahead forecasts with relatively more probability mass at wider intervals of NIM, suggesting a relatively more risk-averse aggregate forecast. The QR model produces narrower aggregate density forecasts relative to the FE-OLS and shows greater skewness in the density forecasts, for example, in the 2007-2009 GFC period and again in the post-2020 period. The FE-QAR model model produces the narrowest aggregate density forecasts across all specifications, and shows relatively larger amounts of skewness in the lower tails of the aggregate distribution until the out-of-sample period, in which both upside and downside tail risks increase.

The BPQR-AL model appears to blend some characteristics of the FE-OLS, QR, and FE-QAR models by including individual effects in a quantile regression framework. The aggregate BPQR-AL density forecasts have relatively larger tails compared to the QR model but place lower probability mass on the tails relative to the FE-OLS model, suggesting more forecast uncertainty relative to the QR model and lower probability of tail events relative to the FE-OLS model. Relative to the FE-QAR model, the BPQR-AL approach has larger tails at both ends of the aggregate distribution, likely driven by the adaptive LASSO shrinkage that places less shrinkage on unobserved individual heterogeneity, which explains more variation near the tails of the individual bank distributions. Additionally, in the post-2020 period, the aggregate BPQR-AL density forecasts do a better job of capturing risks associated with the more extreme tail movements of NIM relative to the FE-OLS, QR, and FE-QAR models.

8 Robustness Checks

I perform additional robustness checks on the BPQR-AL model specification and resulting forecasts. First, I examine the sensitivity of the BPQR-AL model results when excluding macroeconomic variables from the set of covariates. Next, I examine the sensitivity of the forecast results to a specification that uses an informative Minnesota-style prior specification. Last, I compare forecast accuracy using a specification that allows for greater flexibility based on the size subsample results shown in Section 6.4. In each case, I compare results using the a ratio of the CRPS and quantile-weighted CRPS of competing models, and also include a one-sided Diebold-Mariano test to quantify the accuracy of the density forecasts and the statistical improvement relative to alternative BPQR-AL specifications.

8.1 Importance of Macroeconomic Factors

The individual density forecasts produced by the BPQR-AL model are conditioned on four autoregressive terms, macroeconomic factors, and a set of controls to deal with observable and unobservable individual heterogeneity. It is useful to understand the importance of including the macroeconomic interest rate and credit risk factors in the model, given the relative magnitude of the summed autoregressive terms. If the macroeconomic factors are not economically significant for model predictions, then a model design that excludes the covariate vector Z in equation (1) may show similar or perhaps improved forecast performance. For comparison, I estimate the BPQR-AL model of bank NIM on a set of covariates that excludes macroeconomic interest rate and credit risk factors but retains the remaining balance sheet and persistence covariates for comparison against the full model.

Table 4 reports ratios of the CRPS and quantile-weighted CRPS results averaged across banks and time during the in-sample and pseudo out-of-sample periods across model specifications, and includes results of a one-sided DM test where the alternative hypothesis is that the BPQR-AL model that includes macroeconomic factors is statistically more accurate than the model that excludes them. The results show that a model that includes the macroeconomic factors, both in-sample and out-of-sample, outperforms a model that excludes those factors across all score metrics. The DM test results also show that the models with macroeconomic factors are statistically more accurate than the model without macroeconomic factors. Given that credit spreads remained relatively benign during the post-2020 period while interest rate movements were relatively large, this offers additional evidence that distributional sensitivity to interest rates matters for bank NIMs.

8.2 Prior Sensitivity

Another dependency of the individual bank density forecasts produced by the model are the prior specifications necessitated by the Bayesian framework, which were set to be uninformative for the covariate parameter vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, and scale parameter σ given the the relatively large dimension of panel data available. Carriero, Clark, and Marcellino (2024) show that Bayesian shrinkage can improve forecast accuracy in quantile regression estimation of tail risks in cases where sample size may be small. Despite my large sample size afforded by the panel setting, I check the sensitivity of the results to specifying an informative prior specification on the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ in the hierarchical specification of (7). To quantify the sensitivity, I compare the density forecast performance of the model with uninformative priors relative to a model with informative Minnesota-style priors. I follow Carriero, Clark, and Marcellino (2024) and Carriero, Clark, and Marcellino (2020) in specifying a Minnesota-style prior specification for the BPQR-AL model, which shrinks the lagged dependent variable coefficients by less than the other variable coefficients while accounting for the relative standard deviations of independent variable j and the dependent variable, σ_j and σ_{nim}, respectively.

In all cases, the prior mean vectors $\mathbf{b_0}$ and $\mathbf{g_0}$ remain at zero. For the intercept, the prior standard deviation remains uninformative, set to

$$B_0^{intercept} = 1000\sigma_{nim}$$

For each variable j in vectors $\mathbf{x_{it}}$ and $\mathbf{z_{it}}$, prior specifications are set to

$$B_0^j = \lambda_1 \lambda_2 \frac{\sigma_{nim}}{\sigma_j}$$

For each lag l of the dependent variable, prior specifications are set to

$$B_0{}^l = \frac{\lambda_1}{l}$$

I estimate σ_{nim} and σ_j using the sample variances in the data as of each forecast date.[16] The hyperparameters λ_1 and λ_2 are set equal to 0.2, following Carriero, Clark, and Marcellino (2020).

Table 5 reports the ratios of the CRPS and quantile-weighted CRPS results averaged across banks and time during the in-sample and pseudo out-of-sample periods for BPQR-AL model specifications with uninformative priors relative to the Minnesota-style priors. The table includes results of a one-sided DM test where the alternative hypothesis is that the BPQR-AL model with uninformative priors is statistically less accurate than the model with informative priors. The results show that a model with uninformative priors is not meaningfully or statistically less accurate than the Minnesota-style prior model across the in-sample and out-of-sample periods.

8.3 Bank Size Subsample Models

Parameter results shown in Section 6.4 suggest the possibility of parameter estimate heterogeneity across the interest rate factors included in the model. Results shown for density forecasts thus far have relied on estimation across the full sample of banks, but additional flexibility could be added to the BPQR-AL model forecasts when estimating parameters

[16] For the in-sample period, the standard deviation is based on the entire in-sample period. Each subsequent standard deviation in the pseudo out-of-sample period is based on each successive update of the model covariates.

by the bank size groupings previously used, which group banks into size categories of less than $10 billion, $10 billion to $100 billion, and greater than $100 billion in real total assets, and then comparing the forecast accuracy across the size-subsample models to the pooled model to examine if there are improvements in forecast accuracy. If parameter heterogeneity across bank size groups is important, then the additional parameter flexibility of using the size-subsample models could result in more accurate forecasts relative to the full sample approach.

Table 6 reports ratios of the CRPS and quantile-weighted CRPS results averaged across banks and time during the in-sample and pseudo out-of-sample periods for BPQR-AL model estimations based on the full sample estimation relative to the size subsample estimations. The table includes results of a one-sided DM test where the alternative hypothesis is that the BPQR-AL model based on size subsamples is statistically more accurate than the model based on the full sample. The CRPS results suggest that the size subsample model density forecasts are statistically more accurate during the in-sample period. Perhaps unexpectedly, these results do not hold during the pseudo out-of-sample period, where the size subsample model is statistically equivalent to the full sample model. This could be a result of the full sample BPQR-AL approach sufficiently capturing the size-based heterogeneity, given that a measure of bank size is already included as a control variable in the model. The results show that a model with uninformative priors is not meaningfully or statistically less accurate than the Minnesota-style prior model across the in-sample and out-of-sample periods.

9 Measures of Tail Risk

At a micro prudential level, probabilistic measures of risk can be useful to banking supervisors for identifying individual banks that may be exposed to relatively more adverse tail risks. The BPQR-AL model allows for construction of density forecasts for each bank in the sample at each time period. Figure 19 plots the estimated individual bank skewed-t density forecasts from BPQR-AL pseudo out-of-sample year-end estimations from 2019 Q4-2022 Q4, representing the final in-sample forecast and the subsequent year-end pseudo out-of-sample forecasts. From 2019 Q4-2021 Q4, many of the bank-level density forecasts show a leftward location shift and increased dispersion, while some banks show increasing uncertainty as interest rates changed and banks adjusted their balance sheets. By 2022 Q4, bank density forecasts moved rightward as short-term interest rates increased rapidly. These plots suggest that density forecasts of the BPQR-AL model produce variation over time and across banks for the predicted location, scale and shape of the densities.

One way to measure tail behavior of the conditional density forecasts is the expected shortfall and longrise approach. For a chosen probability π, the expected NIM shortfall SF for bank i at forecast horizon $t+1$ is defined as

$$SF_{i,\ t+1}^{NIM} = \frac{1}{\pi} \int_0^{\pi} \hat{F}^{-1}_{y_{i,\ t+1}|\ X_t,\ Z_{t+1}}(\tau | X_t,\ Z_{t+1}) d\tau$$

and expected NIM longrise LR is defined as

$$LR_{i,\ t+1}^{NIM} = \frac{1}{\pi} \int_{1-\pi}^{1} \hat{F}^{-1}_{y_{i,\ t+1}|\ X_t,\ Z_{t+1}}(\tau | X_t,\ Z_{t+1}) d\tau$$

where $\hat{F}^{-1}_{y_{i,\ t+1}|\ X_t,\ Z_{t+1}}$ is the inverse conditional CDF forecast obtained for bank i from fitting the BPQR-AL quantile forecasts to the skewed-t distribution.

I compute expected longrise and shortfall at the 5 percent level by setting $\pi = 0.05$. To impose economic context on the longrise and shortfall measures, I compute the measures as a share of bank average total assets so that they may be interpreted as the change in bank's bottom-line profitability, return on average assets (ROA), under tail outcomes. ROA is defined as the total net income of a bank divided by its average total assets. Net income is important for bank regulatory capital because retained earnings is a primary way that banks can build regulatory capital levels, suggesting that it is important for financial stability more broadly. NIM is one component of ROA, along with noninterest income and expenses, provisions, and other sources of net income. In this context, the expected shortfall (longrise) can be interpreted as the estimated conditional average loss to a bank's total profits in extreme NIM tail events, which would then have implications for a bank's ability to build capital via retained earnings. I follow Giglio et al. (2021) in computing the cross-sectional dispersion of these risk measures across banks over time.

Figure 20 shows the distribution of ROA for bank holding companies in the sample from 1998-2022. The plot shows the 1st, 5th, 25th, 50th, 75th, and 99th percentiles of the distribution. In benign periods, average ROA is roughly 1.5 percent of average total assets. During and shortly after the GFC period, average ROA decreased to 0.25 percent of average total assets and extended to less than negative 10 percent of total assets in the extreme tails of the distribution. In the more recent post-2020 period, bank ROA shows the largest decline in the extreme tails of the distribution since the GFC period as a result of stressed bank NIM and increased provision expenses.[17]

Figure 21 shows the 1st, 5th, 25th, 50th, 75th, and 99th percentiles of the cross-sectional distribution of 5 percent expected longrise and shortfall changes in bank ROA from extreme NIM tail events for all banks in the sample. In both cases, the extreme 1st and 99th tails of the cross-sectional expected longrise and shortfall distribution show variation across economic fluctuations. The longrise measure shows that there is a large and persistent amount of upside potential for increased ROA across the distribution prior to and during the GFC period, which subsequently decreased, and has since normalized back to pre-2007 levels. The shortfall measure shows relatively more volatility at the extreme lower tails of the cross-sectional distribution, particularly following recessionary periods. The magnitude of ROA downside risks suggested by the expected shortfall measure peaks early in the sample period and again in the GFC period but declines gradually afterward until a notable increase in downside risk beginning in 2021. The size of the upside and downside risks are substantial relative to bank ROA across the distribution. The shortfall is sizable enough to move a substantial number of banks near or into negative ROA levels, suggesting that tail risks to bank NIM are material to overall bank profitability.

10 Conclusion

Bank net interest margins are an important component of bank profits, which have been historically stable, on average, and are used to build bank capital so that they matter for

[17] See https://www.federalreserve.gov/publications/2020-november-supervision-and-regulation-report-banking-system-conditions.htm.

overall financial stability. The post-2020 period has shown a break in the historical stability of bank net interest margins against the backdrop of relatively large changes in interest rates, and the cross-sectional distribution suggests the possibility of increased tail risks. This was at least partially a result of interest income declining by more than interest expenses due to zero lower bound constraints. Increased volatility in the tails of the bank NIM distribution suggests that distributional heterogeneity is important to consider when attempting to predict bank net interest margins, and tail risks are increasingly important for regulators and policymakers.

I quantify differential tail sensitivities of bank net interest margins to interest rates, credit risk and own persistence in a quantile regression setting. I follow the general frameworks of Covas, Rump, and Zakrajšek (2014) and Giglio et al. (2021), and examine a relatively large sample of US bank holding companies with at least /$500 million in total assets from 1998-2022. I utilize a novel Bayesian dynamic panel quantile regression approach in the empirical analysis that deals with the problem of incidental parameter bias when applying nonlinear estimation to panel data settings. This approach accounts for observed and unobserved heterogeneity, and heterogeneity in sensitivities to model covariates, as well as providing a natural framework to recover the estimated conditional distribution of bank NIM for individual banks and aggregate outcomes.

I find evidence of heterogeneity in bank sensitivities to interest rate and credit risk in the cross-section. I also find evidence that banks NIM sensitivity to interest rates has increased since 2020, and show that these increased sensitivities are driven by higher interest rate sensitivities for loan interest income and demand deposit interest expense. My results suggest that sensitivity to interest rates increased by a relatively larger margin at the conditional tails of the distribution relative to the median, and specifically for relatively larger banks. The interest rate tail sensitivities that I estimate, relative to the median estimates, are economically meaningful for bank NIM when interest rate movements are large, such as those observed in the post-2020 period.

Lastly, I compare density forecasts implied by the BPQR-AL model against fixed-effect OLS, pooled quantile regression, and the FE-QAR model framework used by Covas, Rump, and Zakrajšek (2014). I show that the forecasts produced by the BPQR-AL framework statistically outperform the competitor models in a pseudo out-of-sample framework by incorporating quantile-varying individual effects with adaptive LASSO shrinkage and quantile-varying sensitivities to covariates. These results suggest that the heterogeneity estimated by the quantile regressions with individual effects should be considered when evaluating the sensitivity of bank NIM to macroeconomic outcomes. Robustness checks suggest that the model forecasts are robust to prior specification and estimation sample choices, and that macroeconomic factors matter for bank NIM forecast accuracy. Density forecasts also allow for computation of expected shortfall and longrise measures, and these measures show that tail behavior of bank NIM has material implications for risk to bottom-line profitability of banks.

While suggestive, the results in this paper are not meant to be interpreted as causal given the reduced-form approach of the model specification. Recent work has carefully explored channels that drive heterogeneity in bank NIM (for example, Williams (2020)). Future research could extend this analysis by offering a more complete analysis of the balance sheet channels that drive heterogeneity across the conditional distribution in a quantile regression context. A structural approach to describing sources of increased heterogeneity in bank NIM in the post-2020 period would be a useful avenue in future work for generating testable model predictions.

11 References

Abdymomunov, Azamat, Jeffery Gerlach, and Yuji Sakurai. 2023. "Interest Rate Risk in the U.S. Banking Sector." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4395529.

Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone. 2019. "Vulnerable Growth." *American Economic Review* 109 (4): 1263–89. https://doi.org/10.1257/aer.20161923.

Aghamohammadi, A., and S. Mohammadi. 2017. "Bayesian Analysis of Penalized Quantile Regression for Longitudinal Data." *Statistical Papers* 58 (4): 1035–53. https://doi.org/10.1007/s00362-015-0737-4.

Alessandri, Piergiorgio, and Benjamin D. Nelson. 2015. "Simple Banking: Profitability and the Yield Curve." *Journal of Money, Credit and Banking* 47 (1): 143–75. https://doi.org/10.1111/jmcb.12172.

Allen, Linda. 1988. "The Determinants of Bank Interest Margins: A Note." *Journal of Financial and Quantitative Analysis* 23 (2): 231–35. https://doi.org/10.2307/2330883.

Angbazo, Lazarus. 1997. "Commercial Bank Net Interest Margins, Default Risk, Interest-Rate Risk, and Off-Balance Sheet Banking." *Journal of Banking & Finance* 21 (1): 55–87. https://doi.org/10.1016/S0378-4266(96)00025-8.

Azzalini, Adelchi, and Antonella Capitanio. 2003. "Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t -Distribution." *Journal of the Royal Statistical Society Series B: Statistical Methodology* 65 (2): 367–89. https://doi.org/10.1111/1467-9868.00391.

Begenau, Juliane, and Erik Stafford. 2022. "Unstable Inference from Banks' Stable Net Interest Margins." *Available at SSRN 4136866*. https://doi.org/10.2139/ssrn.4136866.

Borio, Claudio, Leonardo Gambacorta, and Boris Hofmann. 2017. "The Influence of Monetary Policy on Bank Profitability." *International Finance* 20 (1): 48–63. https://doi.org/10.1111/infi.12104.

Busetti, Fabio. 2017. "Quantile Aggregation of Density Forecasts." *Oxford Bulletin of Economics and Statistics* 79 (4): 495–512. https://doi.org/10.1111/obes.12163.

Carriero, Andrea, Todd E Clark, and Massimiliano Marcellino. 2020. "Nowcasting Tail Risks to Economic Activity with Many Indicators." Federal Reserve Bank of Cleveland.

https://doi.org/10.26509/frbc-wp-202013r2.

———. 2024. “Specification Choices in Quantile Regression for Empirical Macroeconomics.” *Journal of Applied Econometrics*. https://doi.org/10.1002/jae.3099.

Covas, Francisco, Ben Rump, and Egon Zakrajšek. 2014. “Stress-Testing US Bank Holding Companies: A Dynamic Panel Quantile Regression Approach.” *International Journal of Forecasting* 30 (3): 691–713. https://doi.org/10.1016/j.ijforecast.2013.11.003.

Di Lucido, Katherine, Anna Kovner, and Samantha Zeller. 2017. “Low Interest Rates and Bank Profits.” *Federal Reserve Bank of New York Liberty Street Economics (Blog), June 21, 2017*. https://libertystreeteconomics.newyorkfed.org/2017/06/low-interest-rates-and-bank-profits/.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl. 2021. “Banking on Deposits: Maturity Transformation Without Interest Rate Risk.” *The Journal of Finance* 76 (3): 1091–1143. https://doi.org/10.1111/jofi.13013.

English, William. 2002. “Interest Rate Risk and Bank Net Interest Margins.” *BIS Quarterly Review*. https://econpapers.repec.org/article/bisbisqtr/0212g.htm.

English, William, Skander J. Van Den Heuvel, and Egon Zakrajšek. 2018. “Interest Rate Risk and Bank Equity Valuations.” *Journal of Monetary Economics* 98 (October): 80–97. https://doi.org/10.1016/j.jmoneco.2018.04.010.

Ennis, Huberto M, and Arantxa Jarque. 2021. “Bank Lending in the Time of COVID.”

Flannery, Mark. 1981. “Market Interest Rates and Commercial Bank Profitability: An Empirical Investigation.” *The Journal of Finance* 36 (5): 1085–1101. https://doi.org/10.1111/j.1540-6261.1981.tb01078.x.

Flannery, Mark, and Sorin M. Sorescu. 2023. “Partial Effects of Fed Tightening on U.S. Banks’ Capital.” *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4424139.

Galvao, Antonio F. 2011. “Quantile Regression for Dynamic Panel Data with Fixed Effects.” *Journal of Econometrics* 164 (1): 142–57. https://doi.org/10.1016/j.jeconom.2011.02.016.

Giglio, Carla, Frances Shaw, Nicolas Syrichas, and Giuseppe Cappelletti. 2021. “Stress-Testing Net Trading Income: The Case of European Banks.” ECB Working Paper 2021/2525. https://papers.ssrn.com/abstract=3797145.

Gneiting, Tilmann, and Matthias Katzfuss. 2014. “Probabilistic Forecasting.” *Annual Review* *of Statistics and Its Application* 1 (1): 125–51. https://doi.org/10.1146/annurev-statistics-062713-085831.

Gneiting, Tilmann, and Adrian E Raftery. 2007. "Strictly Proper Scoring Rules, Prediction, and Estimation." *Journal of the American Statistical Association* 102 (477): 359–78. https://doi.org/10.1198/016214506000001437.

Gneiting, Tilmann, and Roopesh Ranjan. 2011. "Comparing Density Forecasts Using Threshold-and Quantile-Weighted Scoring Rules." *Journal of Business & Economic Statistics* 29 (3): 411–22. https://doi.org/10.1198/jbes.2010.08110.

Granger, Clive WJ, and Ling-Ling Huang. 1997. "Evaluation of Panel Data Models: Some Suggestions from Time Series." https://doi.org/10.2139/ssrn.52400.

Hanweck, Gerald A., and Lisa H. Ryu. 2005. "The Sensitivity of Bank Net Interest Margins and Profitability To Credit, Interest-Rate, and Term-Structure Shocks Across Bank Product Specializations." FDIC Working Paper 05-02. Rochester, NY. https://doi.org/10.2139/ssrn.886727.

Haubrich, Joseph G., and Tristan Young. 2019. "Trends in the Noninterest Income of Banks." *Federal Reserve Bank of Cleveland Economic Commentary*, no. 2019-14. https://doi.org/10.26509/frbc-ec-201914.

Hirtle, Beverly, Anna Kovner, James Vickery, and Meru Bhanot. 2016. "Assessing Financial Stability: The Capital and Loss Assessment Under Stress Scenarios (CLASS) Model." *Journal of Banking & Finance* 69 (August): S35–55. https://doi.org/10.1016/j.jbankfin.2015.09.021.

Ho, Thomas S. Y., and Anthony Saunders. 1981. "The Determinants of Bank Interest Margins: Theory and Empirical Evidence." *The Journal of Financial and Quantitative Analysis* 16 (4): 581. https://doi.org/10.2307/2330377.

Jiang, Erica, Gregor Matvos, Tomasz Piskorski, and Amit Seru. 2023. "Monetary Tightening and U.S. Bank Fragility in 2023: Mark-to-Market Losses and Uninsured Depositor Runs?" https://doi.org/10.1016/j.jfineco.2024.103899.

Koenker, Roger. 2004. "Quantile Regression for Longitudinal Data." *Journal of Multivariate Analysis* 91 (1): 74–89. https://doi.org/10.1016/j.jmva.2004.05.006.

Kozumi, Hideo, and Genya Kobayashi. 2011. "Gibbs Sampling Methods for Bayesian

Quantile Regression." *Journal of Statistical Computation and Simulation* 81 (11): 1565–78. https://doi.org/10.1080/00949655.2010.496117.

Laliberte, Brendan, and Rajdeep Sengupta. 2024. "Why Do Net Interest Margins Behave Differently Across Banks as Interest Rates Rise? Evidence from the Recent Tightening Cycle." *Federal Reserve Bank of Kansas City, Economic Review*. https://doi.org/10.18651/ER/v109n1LaliberteSengupta.

McShane, R. W., and I. G. Sharpe. 1985. "A Time Series/Cross Section Analysis of the Determinants of Australian Trading Bank Loan/Deposit Interest Margins: 1962-1981." *Journal of Banking & Finance* 9 (1): 115–36. https://doi.org/10.1016/0378-4266(85)90065-2.

Nickell, Stephen. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49 (6): 1417–26. https://doi.org/10.2307/1911408.

Park, Trevor, and George Casella. 2008. "The Bayesian Lasso." *Journal of the American Statistical Association* 103 (482): 681–86. https://doi.org/10.1198/016214508000000337.

Samuelson, Paul A. 1945. "The Effect of Interest Rate Increases on the Banking System." *American Economic Review* 35 (1): 16–27. https://www.jstor.org/stable/1810106.

Saunders, Anthony, and Liliana Schumacher. 2000. "The Determinants of Bank Interest Rate Margins: An International Study." *Journal of International Money and Finance* 19 (6): 813–32. https://doi.org/10.1016/S0261-5606(00)00033-4.

Sengupta, Rajdeep, and Fei Xue. 2022. "Do Net Interest Margins for Small and Large Banks Vary Differently with Interest Rates?" *Federal Reserve Bank of Kansas City, Economic Review*. https://doi.org/10.18651/ER/v107n1SenguptaXue.

Tibshirani, Robert. 1996. "Regression Shrinkage and Selection Via the Lasso." *Journal of the Royal Statistical Society: Series B Methodolody* 58 (1): 267–88. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

Williams, Emily. 2020. "Heterogeneity in Net-Interest Income Exposure to Interest Rate Risk and Non-Interest Expense Adjustment."

Yu, Keming, and Rana A. Moyeed. 2001. "Bayesian Quantile Regression." *Statistics & Probability Letters* 54 (4): 437–47. https://doi.org/10.1016/S0167-7152(01)00124-9.

Zarruk, Emilio R, and Jeff Madura. 1992. "Optimal Bank Interest Margin Under Capital

Regulation and Deposit Insurance." *Journal of Financial and Quantitative Analysis* 27 (1): 143–49. https://doi.org/10.2307/2331303.

Zou, Hui. 2006. "The Adaptive Lasso and Its Oracle Properties." *Journal of the American Statistical Association* 101 (476): 1418–29. https://doi.org/10.1198/016214506000000735.

12 Tables

Table 1: Sample includes BHCs with at least $500 million in real total assets from 1998 Q1-2022 Q4. Number of unique BHCs in data: 313. Maturity gap calculation based on English, Van den Heuvel and Zakrajsek (2018). Source: US FRY9-c and Call Report Filings, H.15 interest rate data, ICE Data Indices, LLC

Estimation Sample Summary Statistics, 1998-2022								
Variable	n	Mean	SD	Min	P25	P50	P75	Max
Dependent Variable								
NIM	24326	3.60	0.93	-12.85	3.13	3.60	4.09	10.35
Financial Variables								
3m Tbill	24326	1.82	1.93	0.01	0.11	1.10	3.44	6.20
10y-3m Treasury Spread	24326	1.64	1.14	-0.63	0.68	1.71	2.63	3.61
BBB-10y Spread	24326	1.59	0.83	0.72	1.10	1.42	1.81	5.80
Bank Balance Sheet Variables								
Core Deposits % of Liabilities	24326	69.92	18.04	0.00	62.79	73.85	81.95	98.80
Interest Rate Derivatives % of Assets	24326	5.72	16.60	0.00	0.00	0.87	5.75	760.95
Loans % of IE Assets	24326	70.86	15.65	3.24	64.36	73.69	80.79	175.62
Maturity Gap (years)	24326	4.57	2.17	-2.03	3.01	4.29	5.88	16.31
NCO % of Loans	24326	0.47	1.06	-4.30	0.04	0.19	0.46	25.94
Real Total Assets (USD Billions, 1998=100)	24326	36.07	166.70	0.41	1.22	2.60	7.92	2266.74
Securities % of IE Assets	24326	24.84	13.10	0.00	16.00	22.72	30.95	123.65
Tier 1 Leverage Ratio	24326	9.60	3.89	-2.80	8.01	9.11	10.38	76.26

Table 2: FE-OLS and BPQR-AL estimates, 1998-2022. Total number of bank holding companies in sample: 313. The dependent variable is the level of NIM for bank holding company i at time t. BQRAL posterior means represent the point estimates, and BPQR-AL standard errors are based on 30,000 total draws, of which 10,000 are burn-in draws. FE-OLS (BPQR-AL) 95 percent confidence (credible) intervals in parentheses. FE-OLS standard errors are double-clustered by bank and time. Bank controls include Tier 1 leverage ratio, log of real total assets, loans as a share of interest-earning assets, securities as a share of interest-earning assets, gross notional interest rate derivatives share of assets, net charge-offs as a share of total loans, core deposits as a share of liabilities, and the maturity gap proposed by English, Van den Heuvel and Zakrajsek (2018).

FE-OLS and BPQR-AL Regression Estimates, 1998-2022						
		BPQR-AL Quantiles				
	FE-OLS	0.05	0.25	0.5	0.75	0.95
NIM (4 Lags)	0.812	0.786	0.878	0.901	0.903	0.835
	[0.482,1.141]	[0.749,0.825]	[0.849,0.907]	[0.87,0.933]	[0.875,0.93]	[0.8,0.867]
3m Tbill	0.069	0.048	0.031	0.034	0.046	0.084
	[0.052,0.085]	[0.046,0.05]	[0.029,0.033]	[0.032,0.035]	[0.045,0.048]	[0.081,0.086]
10y-3m Treasury Spread	0.093	0.069	0.043	0.047	0.062	0.094
	[0.07,0.116]	[0.065,0.072]	[0.041,0.046]	[0.045,0.05]	[0.059,0.064]	[0.09,0.098]
BBB-10y Spread	-0.006	-0.04	-0.017	-0.002	0.011	0.043
	[-0.02,0.007]	[-0.043,-0.037]	[-0.019,-0.014]	[-0.005,0]	[0.009,0.013]	[0.039,0.046]
Bank Effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Table 3: Ratio of average continuous ranked probability score (CRPS) and quantile-weighted CRPS across alternative models. Left tail, center, and right tail-weighted CRPS are denoted as CRPS-L, CRPS-C, and CRPS-R, respectively. Model CRPS comparisons against the Bayesian panel quantile regression with adaptive LASSO (BPQR-AL) include (i) fixed-effects OLS (FE-OLS), (ii) pooled quantile regression (QR), and (iii) fixed-effects quantile autoregressive model (FE-QAR). Values less (greater) than 1 indicate the BPQR-AL model produces a smaller (larger) CRPS relative to the comparison model. In-sample period from 1998-2019, pseudo out-of-sample period from 2020-2022. Inference from Diebold–Mariano tests use HAC robust standard errors.

Ratio	Sample	CRPS	CRPS-L	CRPS-C	CRPS-R
BPQR-AL/FE-OLS	In-Sample (1998-2019)	0.86***	0.86***	0.87***	0.85***
BPQR-AL/QR	In-Sample (1998-2019)	1.04	1.05	1.04	1.05
BPQR-AL/FE-QAR	In-Sample (1998-2019)	0.89***	1.02	0.93***	0.75***
BPQR-AL/FE-OLS	Pseudo Out-of-Sample (2020-2022)	0.81***	0.89***	0.8***	0.75***
BPQR-AL/QR	Pseudo Out-of-Sample (2020-2022)	0.9***	0.93***	0.92***	0.84***
BPQR-AL/FE-QAR	Pseudo Out-of-Sample (2020-2022)	0.87***	0.93***	0.92***	0.75***

* denotes statistical significance at the 10% level based on one-sided Diebold–Mariano tests.

** denotes statistical significance at the 5% level based on one-sided Diebold–Mariano tests.

*** denotes statistical significance at the 1% level based on one-sided Diebold–Mariano tests.

Table 4: Ratios of average continuous ranked probability Score comparisons for BPQR-AL model specifications that either include or exclude macroeconomic interest rate and credit risk factors during the in-sample and pseudo out-of-sample periods. Left tail, center, and right tail-weighted CRPS are denoted as CRPS-L, CRPS-C, and CRPS-R, respectively. Values less (greater) than 1 indicate the BPQR-AL model produces a smaller (larger) CRPS relative to the comparison model. In-sample period from 1998-2019, pseudo out-of-sample period from 2020-2022. Inference from Diebold–Mariano tests use HAC robust standard errors.

BPQR-AL CRPS Measures by Inclusion of Macroeconomic Factors					
Ratio	Sample	CRPS	CRPS-L	CRPS-C	CRPS-R
BPQR-AL with/without Macros	In-Sample (1998-2019)	0.95***	0.96***	0.95***	0.93***
BPQR-AL with/without Macros	Pseudo Out-of-Sample (2020-2022)	0.9***	0.92***	0.91***	0.88***

* denotes statistical significance at the 10% level based on one-sided Diebold–Mariano tests.

** denotes statistical significance at the 5% level based on one-sided Diebold–Mariano tests.

*** denotes statistical significance at the 1% level based on one-sided Diebold–Mariano tests.

Table 5: Ratios of average continuous ranked probability score comparisons for BPQR-AL model specifications that use uninformative or Minnesota-style priors during the in-sample and pseudo out-of-sample periods. Left tail, center, and right tail-weighted CRPS are denoted as CRPS-L, CRPS-C, and CRPS-R, respectively. Values less (greater) than 1 indicate the BPQR-AL model produces a smaller (larger) CRPS relative to the comparison model. In-sample period from 1998-2019, pseudo out-of-sample period from 2020-2022. Inference from Diebold–Mariano tests use HAC robust standard errors.

BPQR-AL CRPS Measures by Covariate Prior Specification					
Ratio	Sample	CRPS	CRPS-L	CRPS-C	CRPS-R
Uninformative/Minnesota Priors	In-Sample (1998-2019)	1.00	1.00	1.00	1.00
Uninformative/Minnesota Priors	Pseudo Out-of-Sample (2020-2022)	0.99	1.00	0.99	0.99

* denotes statistical significance at the 10% level based on one-sided Diebold–Mariano tests.

** denotes statistical significance at the 5% level based on one-sided Diebold–Mariano tests.

*** denotes statistical significance at the 1% level based on one-sided Diebold–Mariano tests.

Table 6: Ratios of average continuous ranked probability score comparisons for full and size-subsample BPQR-AL model specifications during the in-sample and pseudo out-of-sample periods. Size subsamples group banks less than $10 billion, $10 billion to $100 billion, and greater than $100 billion for the average of the banks real total assets over the sample period. Left tail, center, and right tail-weighted CRPS are denoted as CRPS-L, CRPS-C, and CRPS-R, respectively. Values less (greater) than 1 indicate the BPQR-AL model produces a smaller (larger) CRPS relative to the comparison model. In-sample period from 1998-2019, pseudo out-of-sample period from 2020-2022. Inference from Diebold–Mariano tests use HAC robust standard errors.

BPQR-AL CRPS Measures for Full Sample Versus Size-Based Subsamples					
Ratio	Sample	CRPS	CRPS-L	CRPS-C	CRPS-R
Full Sample/Size Subsamples	In-Sample (1998-2019)	1.04***	1.05***	1.03***	1.04***
Full Sample/Size Subsamples	Pseudo Out-of-Sample (2020-2022)	0.99	1.00	0.99	0.98

* denotes statistical significance at the 10% level based on one-sided Diebold–Mariano tests.

** denotes statistical significance at the 5% level based on one-sided Diebold–Mariano tests.

*** denotes statistical significance at the 1% level based on one-sided Diebold–Mariano tests.

13 Figures

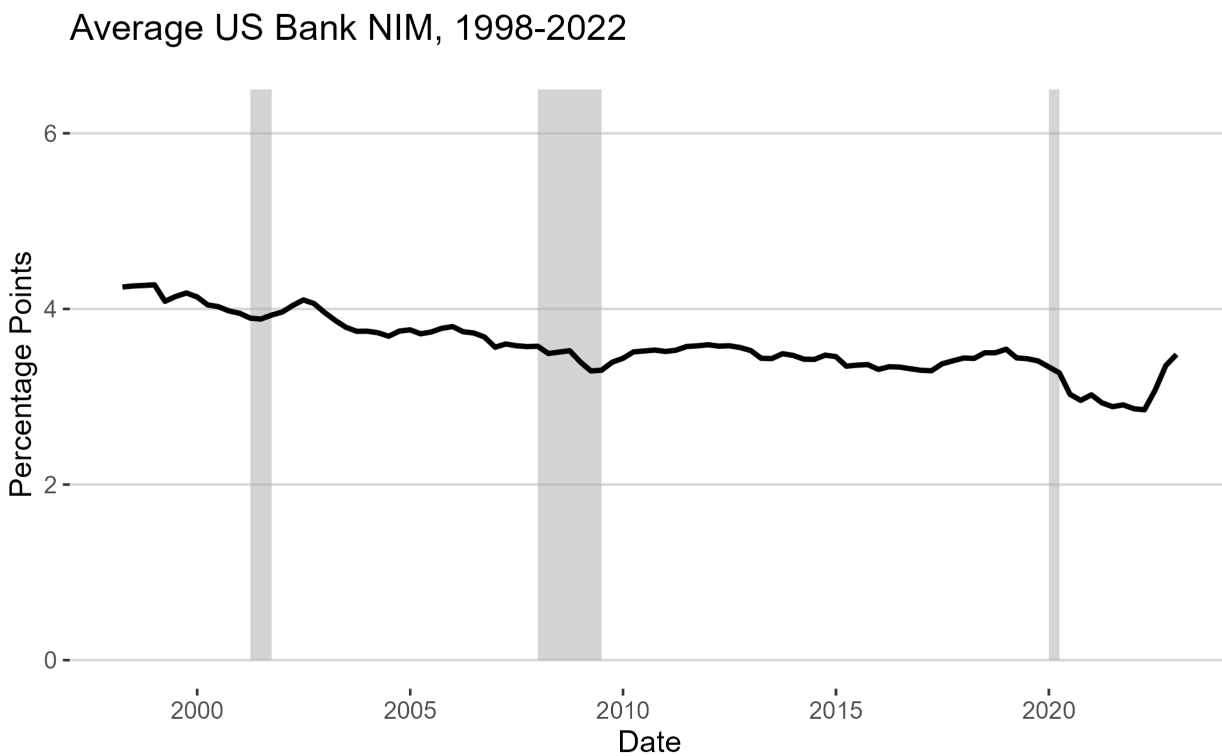

Figure 1: Average annualized bank net interest margin, 1998-2022. Vertical grey areas represent NBER-defined US recessions. Sample excludes bank holding companies with zero balances of total assets, interest-earning assets, or loans. Source: Merger-Adjusted FRY9-c Data (MAY9c)

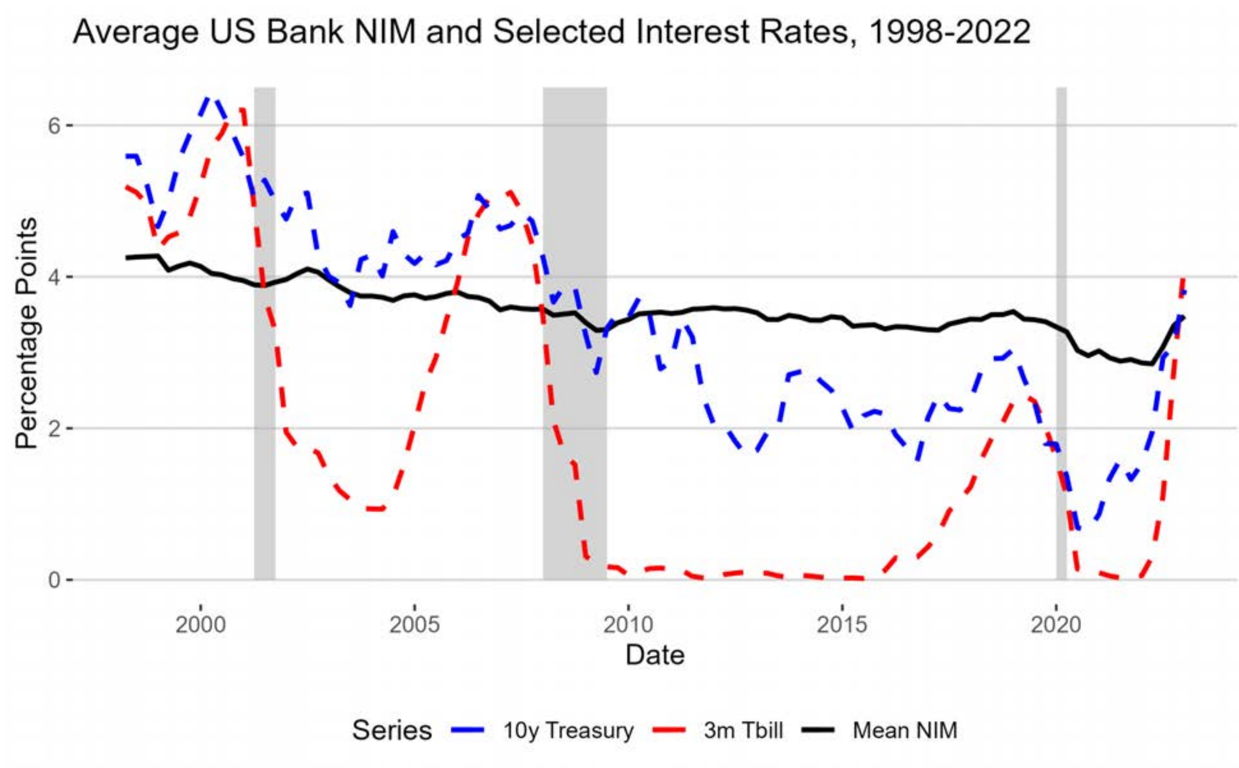

Figure 2: Average annualized bank NIM and selected interest rates, 1998-2022. Vertical grey areas represent NBER-defined US recessions. Sample excludes bank holding companies with zero balances of total assets, interest-earning assets, or loans. Source: Merger-Adjusted FRY9-c Data (MAY9c), H.15 Selected Interest Rates.

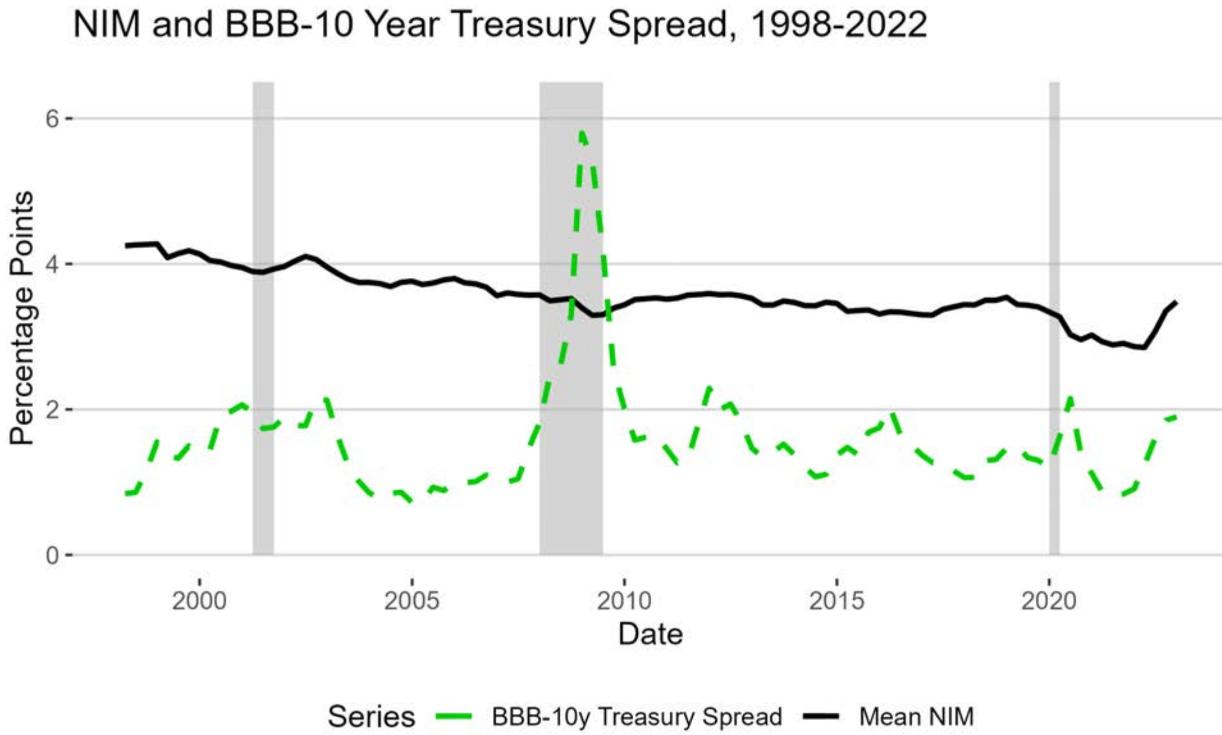

Figure 3: Average annualized bank NIM and BBB-10y spread, 1998-2022. Vertical grey areas represent NBER-defined US recessions. Sample excludes bank holding companies with zero balances of total assets, interest-earning assets, or loans. Vertical grey areas represent NBER-defined US recessions. Source: Merger-Adjusted FRY9-c Data (MAY9c), H.15 Selected Interest Rates, ICE Data Indices, LLC.

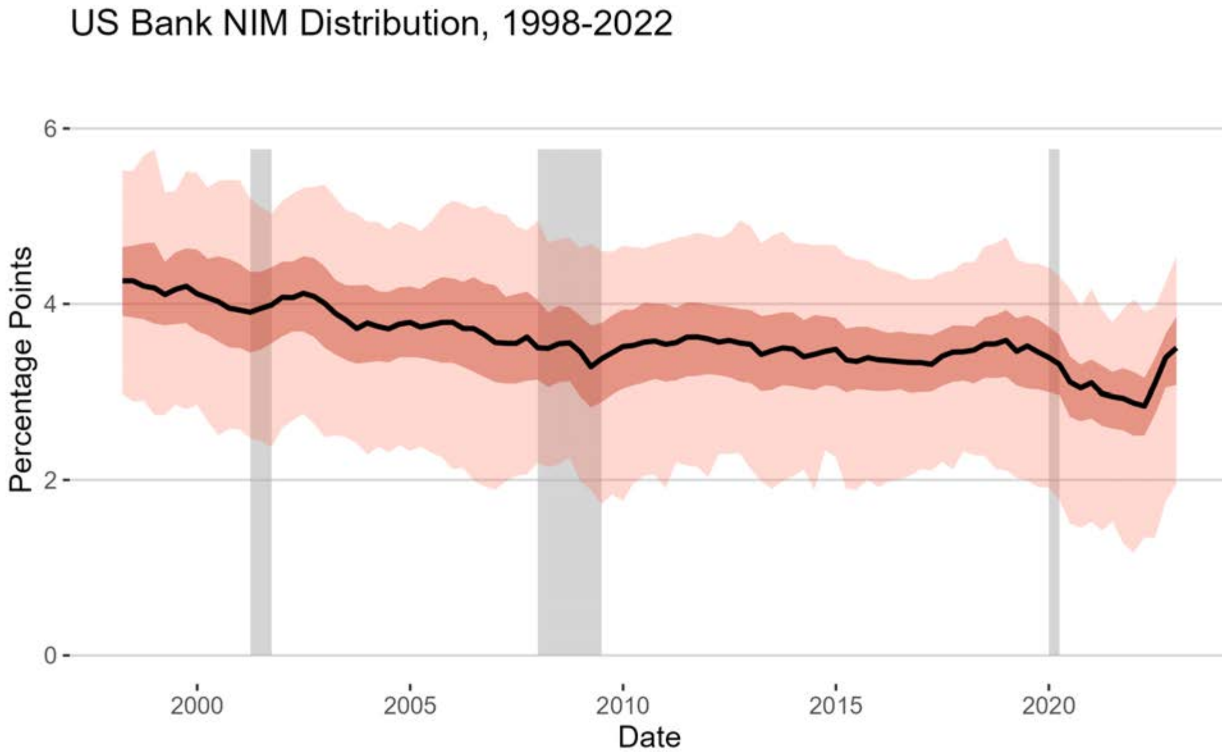

Figure 4: Distribution of Net Interest Margin (annualized), 1998-2022. Shaded areas represent 5th, 25th, 75th, and 95th percentile cross-sectional outcomes, and the solid black line represents median cross-sectional outcome. Sample excludes bank holding companies with zero balances of total assets, interest-earning assets, or loans. Source: Merger-Adjusted FRY9-c Data (MAY9c). Vertical grey areas represent NBER-defined US recessions. Source: US FRY9-c Filings.

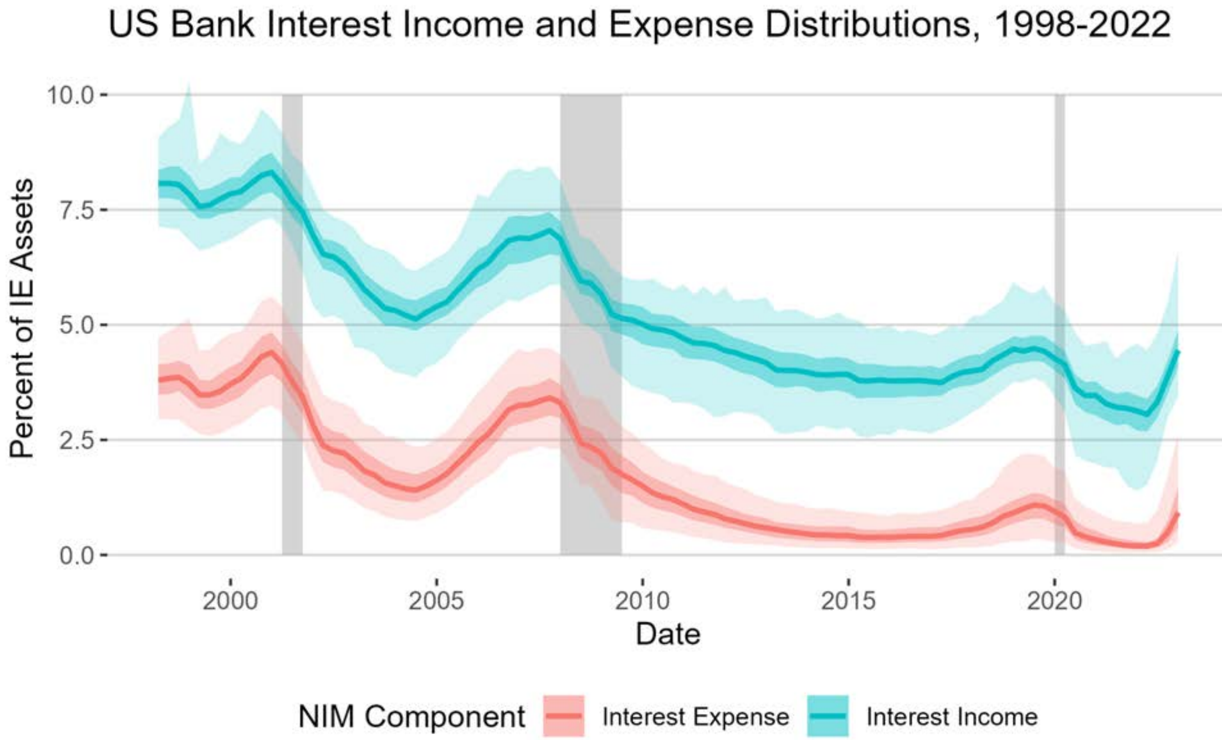

Figure 5: Distribution of Interest Income and Interest Expense as a Percent of Interest-Earning Assets, 1998-2022. Solid line indicates cross-sectional median, dark shaded areas indicate interquartile range, and light shaded areas indicate 5th-95th percentiles. Sample excludes bank holding companies with zero balances of total assets, interest-earning assets, or loans. Vertical grey areas represent NBER-defined US recessions. Source: Merger-Adjusted Y-9c (MAY9c).

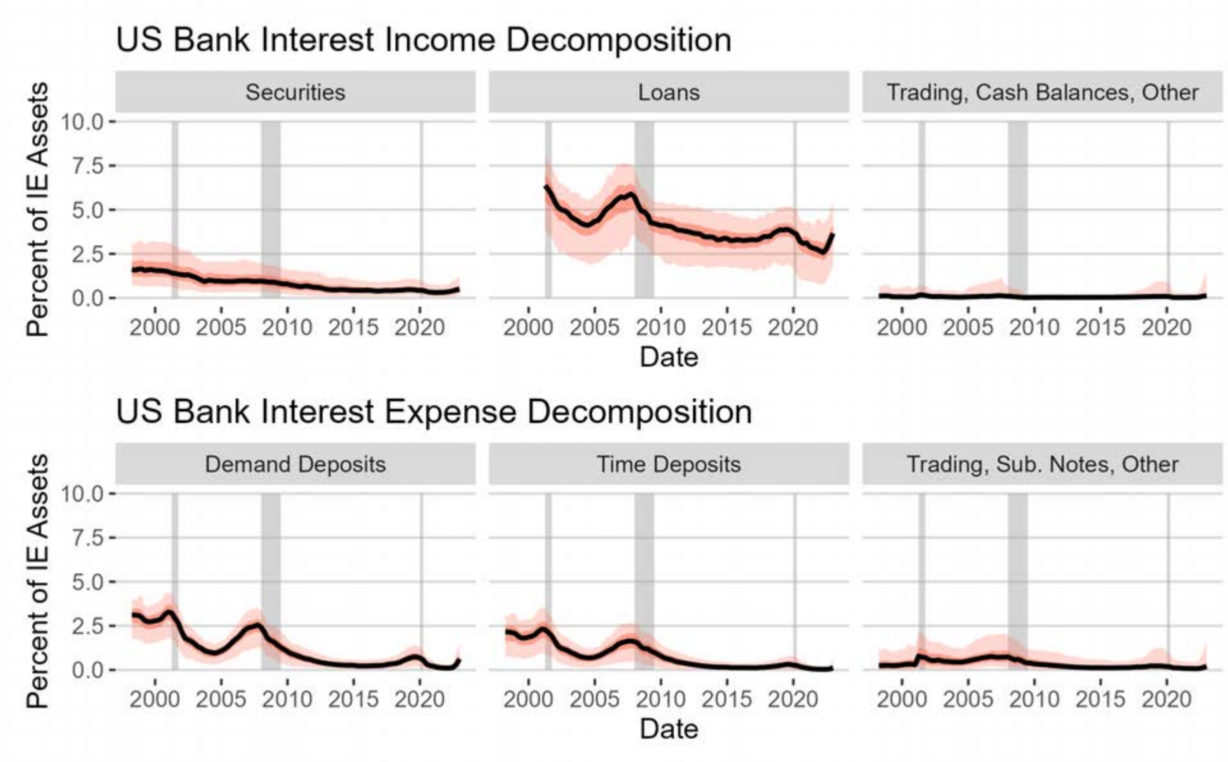

Figure 6: Income and Expense decomposition of US bank net interest margin, 1998-2022. Solid line indicates cross-sectional median, dark shaded areas indicate interquartile range, and light shaded areas indicate 5th-95th percentiles. Vertical grey areas represent NBER-defined US recessions. Source: Merger-Adjusted Y-9c (MAY9c).

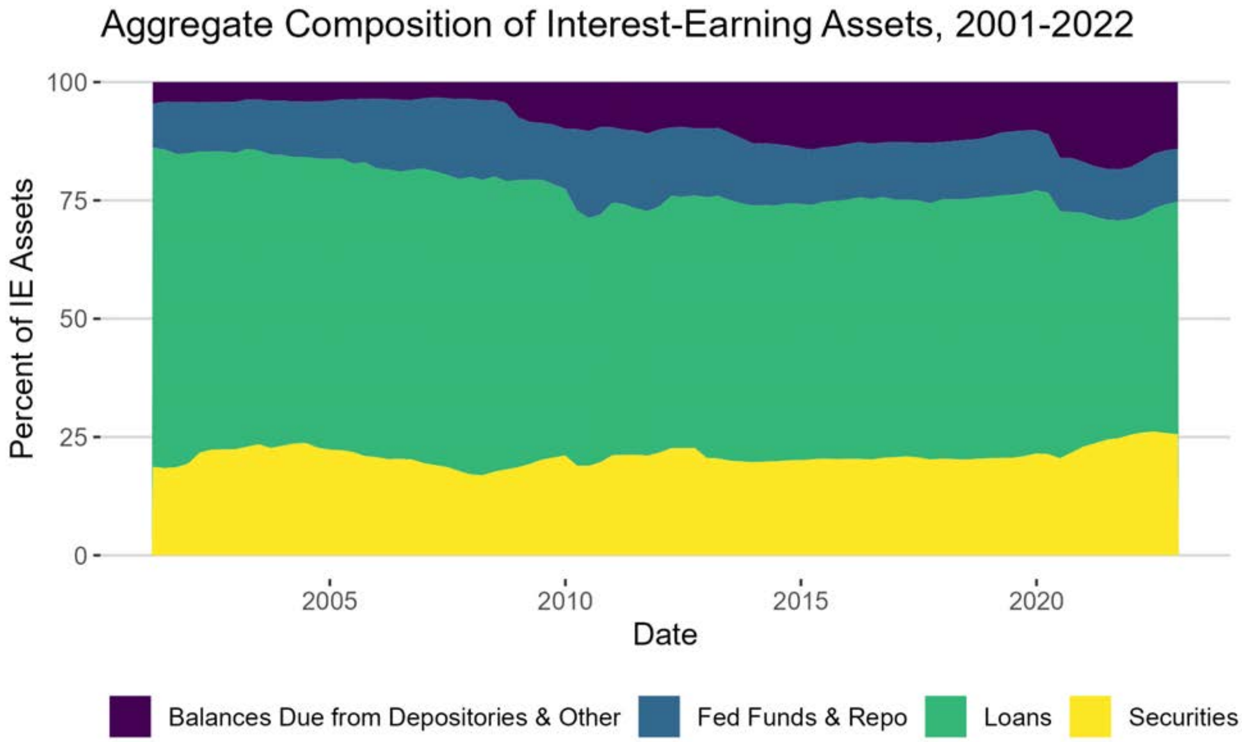

Figure 7: Aggregate components of interest-earning assets as a percent of total interest-earning assets, 2001-2022. 1998-2000 excluded from plot because the "other" category of interest-earning assets was not reported until 2001. Sample excludes bank holding companies with zero balances of total assets, interest-earning assets, or loans. Vertical grey areas represent NBER-defined US recessions. Source: Merger-Adjusted FRY9-c Data (MAY9c)

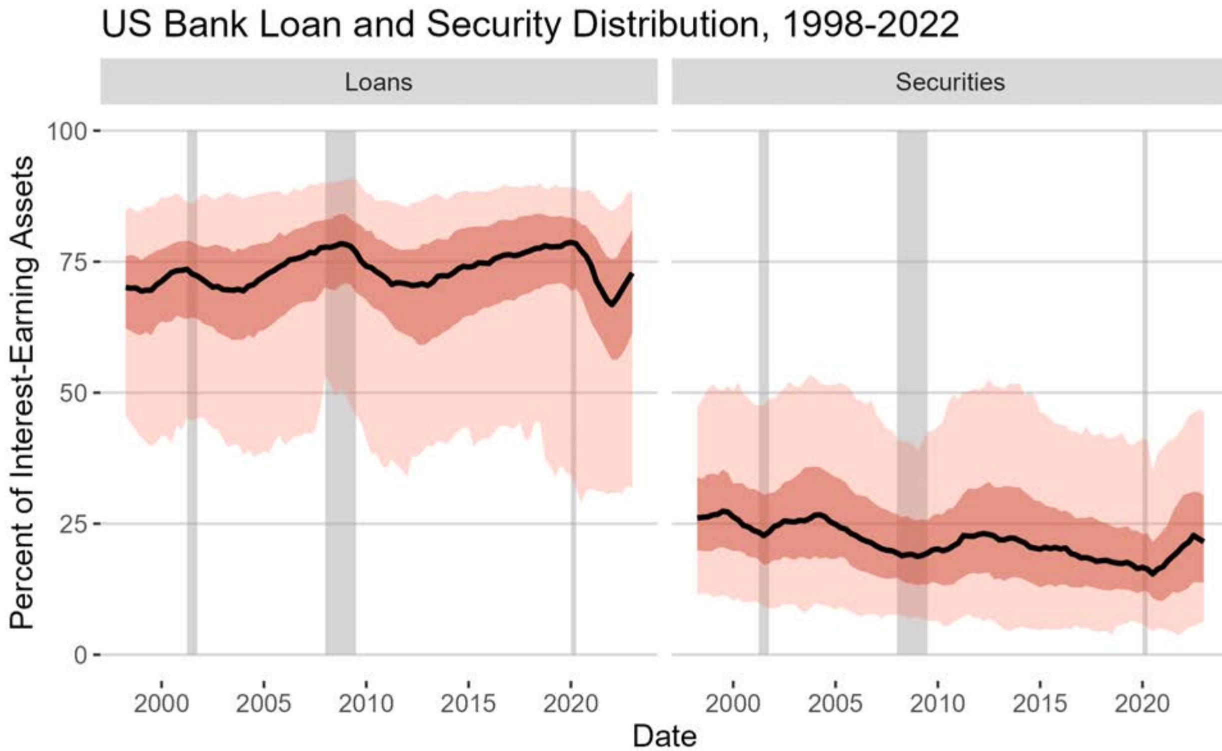

Figure 8: Distribution of loan and security portfolios as a percent of interest-earning assets, 1998-2022. Solid line indicates cross-sectional median, dark shaded areas indicate interquartile range, and light shaded areas indicate 5th-95th percentiles. Sample excludes bank holding companies with zero balances of total assets, interest-earning assets, or loans. Vertical grey areas represent NBER-defined US recessions. Source: Merger-Adjusted Y-9c (MAY9c).

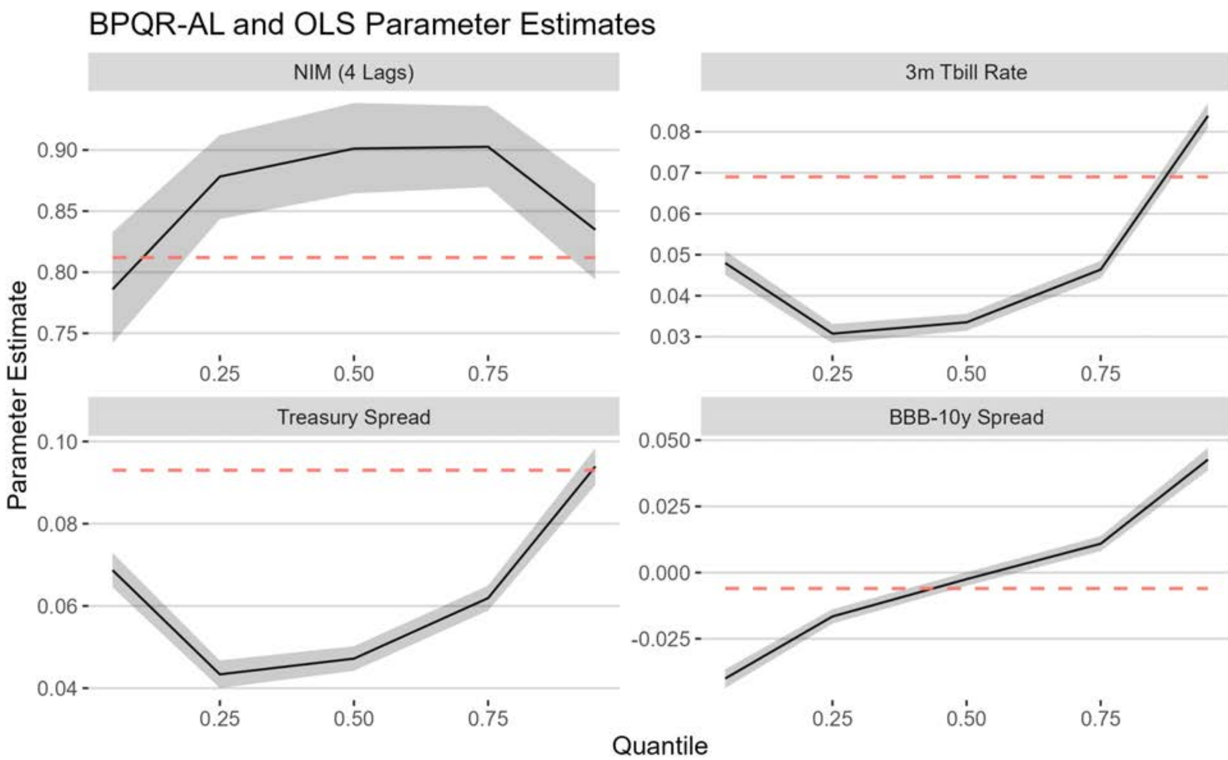

Figure 9: BPQR-AL Quantile Parameter Estimates, 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 35,000 total draws taken from MCMC Gibbs sampling algorithm with 25,000 burn-in draws. Solid black lines represent posterior mean estimates, and shaded areas represent 95% credible intervals. Dashed red line represents FE-OLS estimates. Bank controls in the estimated model include Tier 1 leverage ratio, log of real total assets, loans as a share of interest-earning assets, securities as a share of interest-earning assets, gross notional interest rate derivatives as a share of assets, net charge-offs as a share of total loans, core deposits as a share of liabilities, and the maturity gap proposed by English, Van den Heuvel and Zakrajsek (2018).

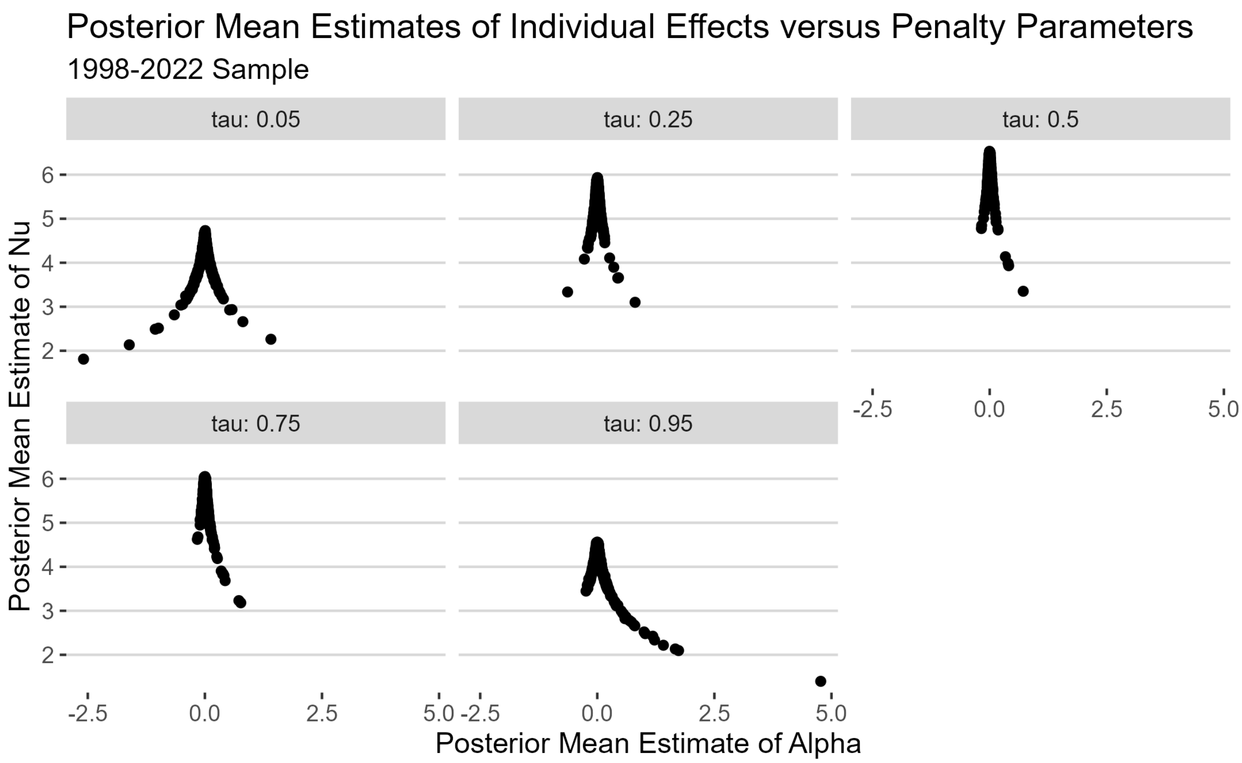

Figure 10: BPQR-AL Individual Random Effects (alpha) Relative to Shrinkage Parameter (nu), 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 35,000 total draws taken from MCMC Gibbs sampling algorithm with 25,000 burn-in draws.

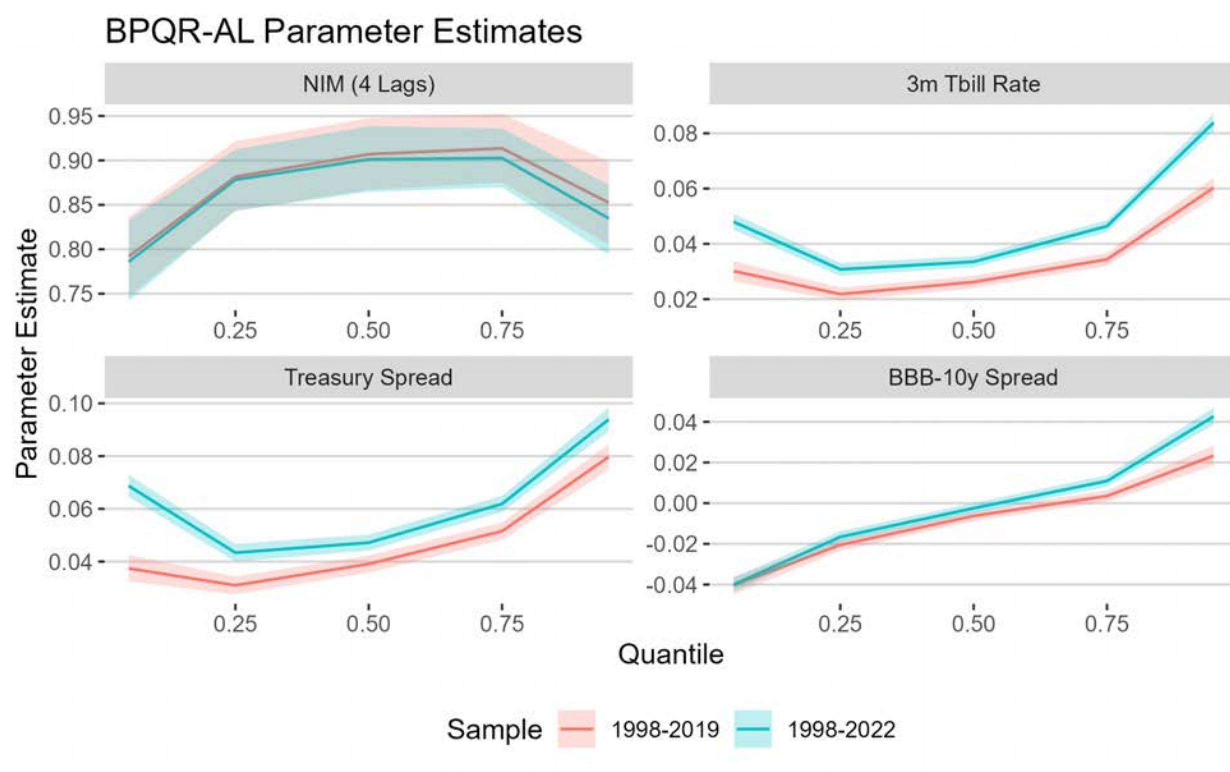

Figure 11: BPQR-AL Quantile Parameter Estimates by Time Period subsamples spanning 1998-2019 and 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 35,000 total draws taken from MCMC Gibbs sampling algorithm with 25,000 burn-in draws. Solid lines represent posterior mean estimates, and shaded areas represent 95% credible intervals.

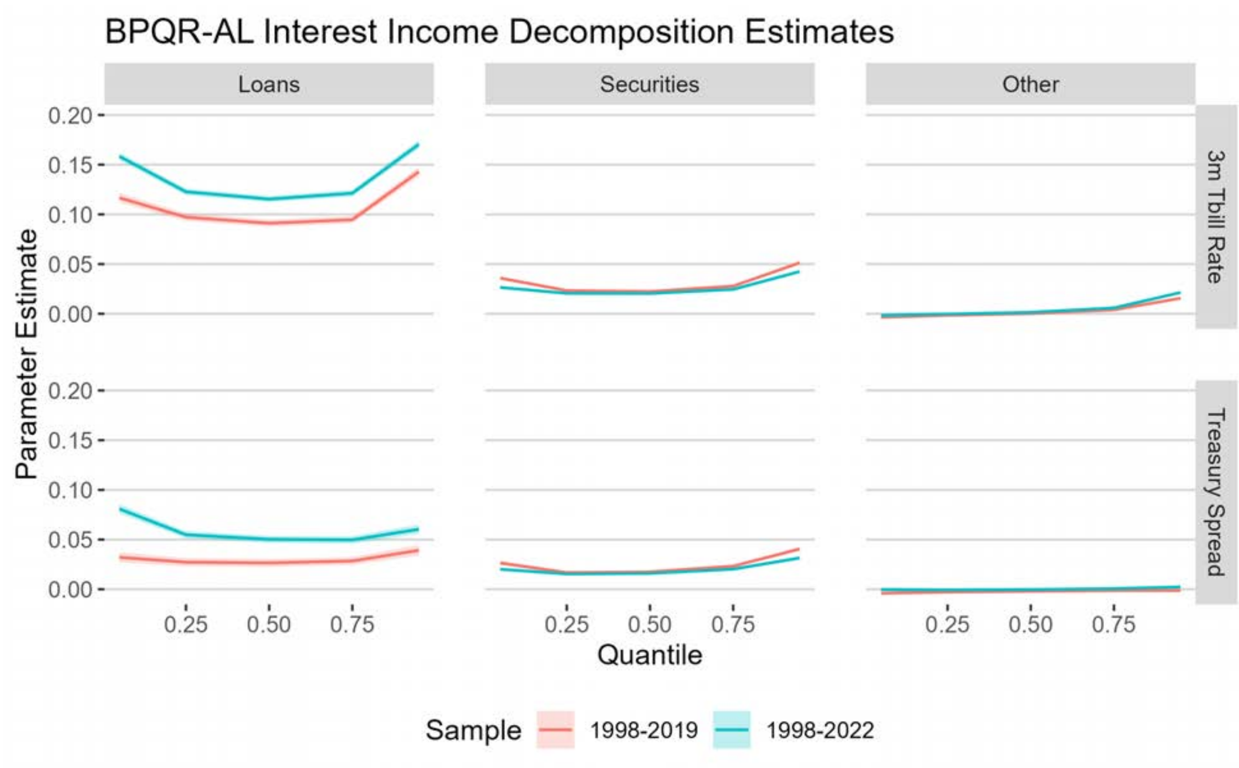

Figure 12: BPQR-AL Quantile Parameter Estimates by Time Period subsamples spanning 1998-2019 and 1998-2022. The 'other' category sources of interest income includes trading, balances due from depository institutions, fed funds and repo, and other. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws. Solid lines represent posterior mean estimates, and shaded areas represent 95% credible intervals.

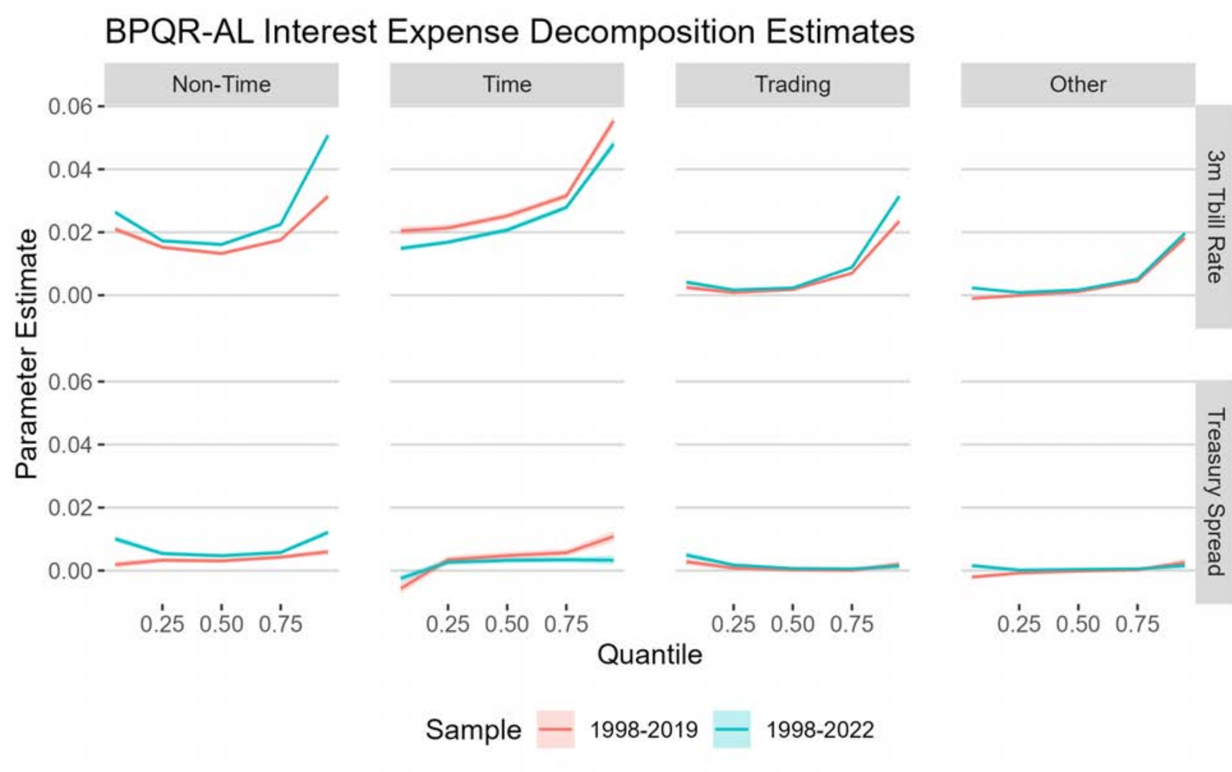

Figure 13: BPQR-AL Quantile Parameter Estimates by Time Period subsamples spanning 1998-2019, and 1998-2022. The 'other' category of interest expenses includes subordinated notes and debentures, fed funds and repo, and other. Measures of interest expense on foreign deposits not included. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws. Solid lines represent posterior mean estimates, and shaded areas represent 95% credible intervals.

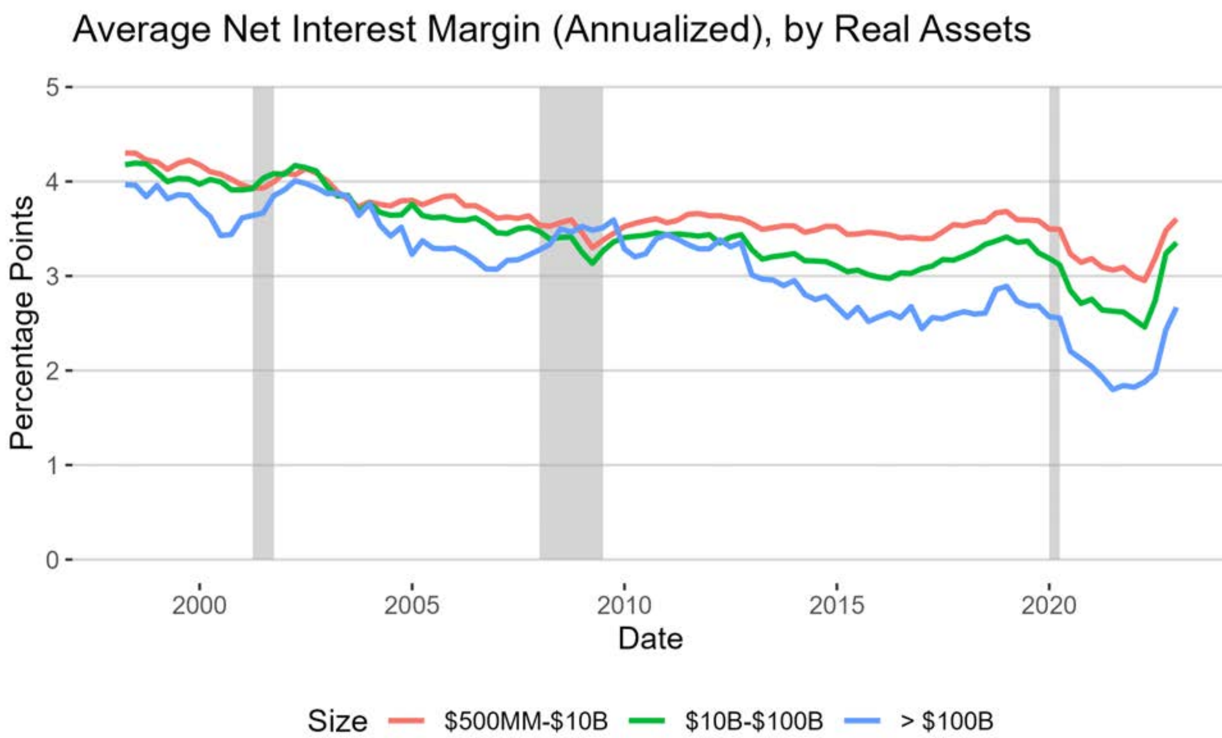

Figure 14: Average Net Interest Margin (Annualized), by Size, 1998-2022. Vertical grey areas represent NBER-defined US recessions. Source: Merger-Adjusted FRY9-c Data (MAY9c)

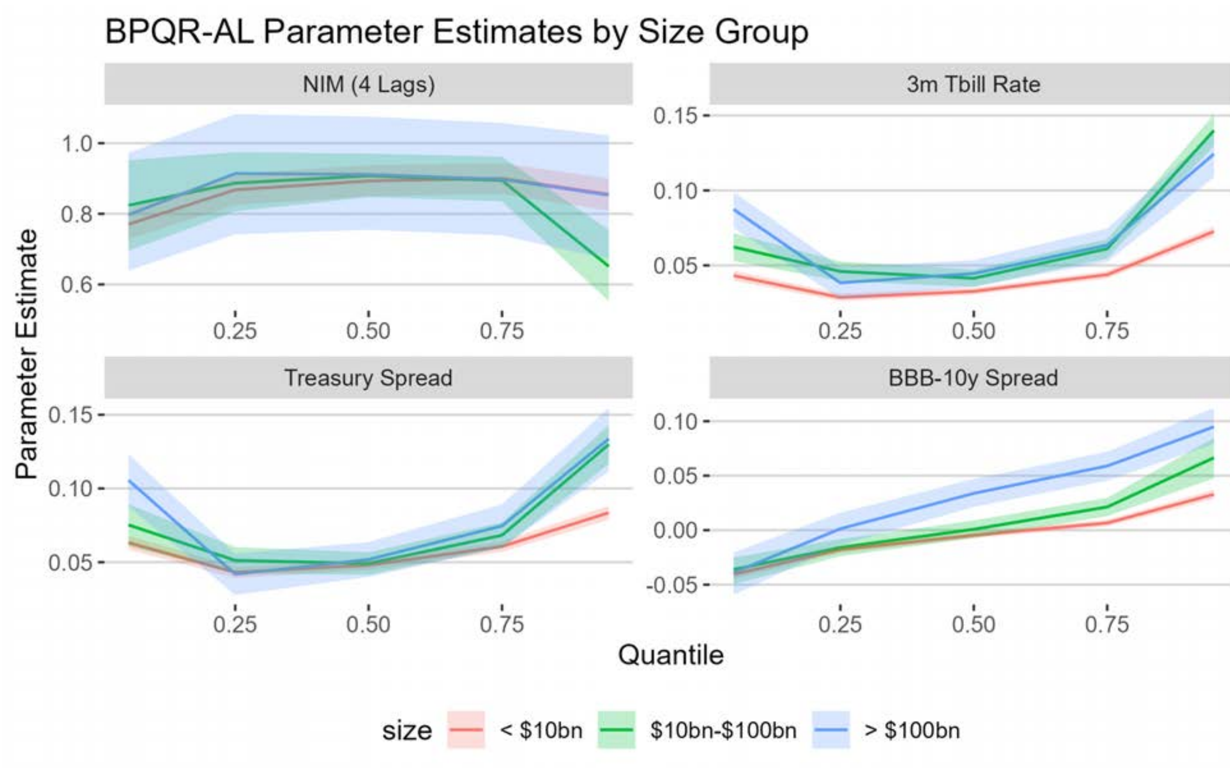

Figure 15: BPQR-AL quantile parameter estimates by bank size group, 1998-2022. Bank size grouping defined by average total assets over the entire sample period. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws. Solid lines represent posterior mean estimates, and shaded areas represent 95% credible intervals.

Figure 16: BPQR-AL Interest rate factor quantile parameter estimates by bank size group and time period. Bank size grouping defined by average total assets over the entire sample period. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws. Solid lines represent posterior mean estimates, and shaded areas represent 95% credible intervals.

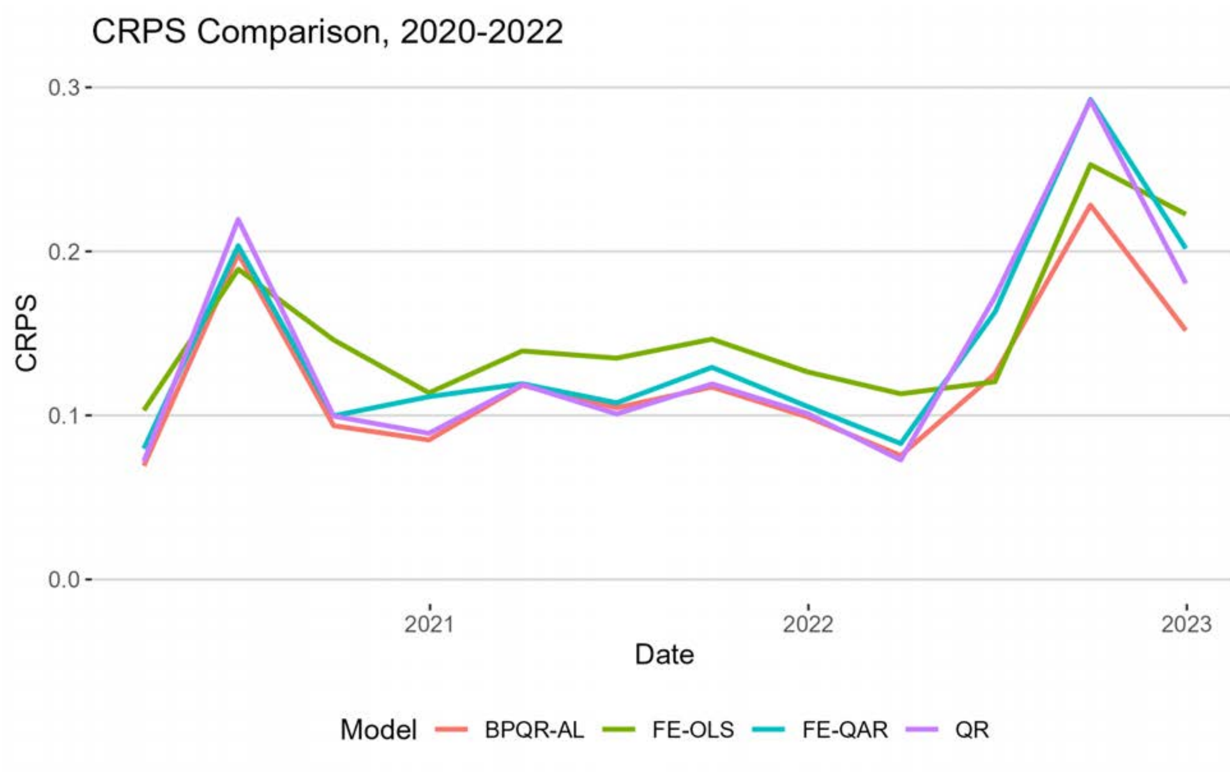

Figure 17: Pseudo out-of-sample CRPS by time period across alternative models. Figure includes fixed-effects OLS (FE-OLS), pooled quantile regression (QR), and Bayesian panel quantile regression with adaptive LASSO (BPQR-AL). Pseudo out-of-sample period spans 2020-2022.

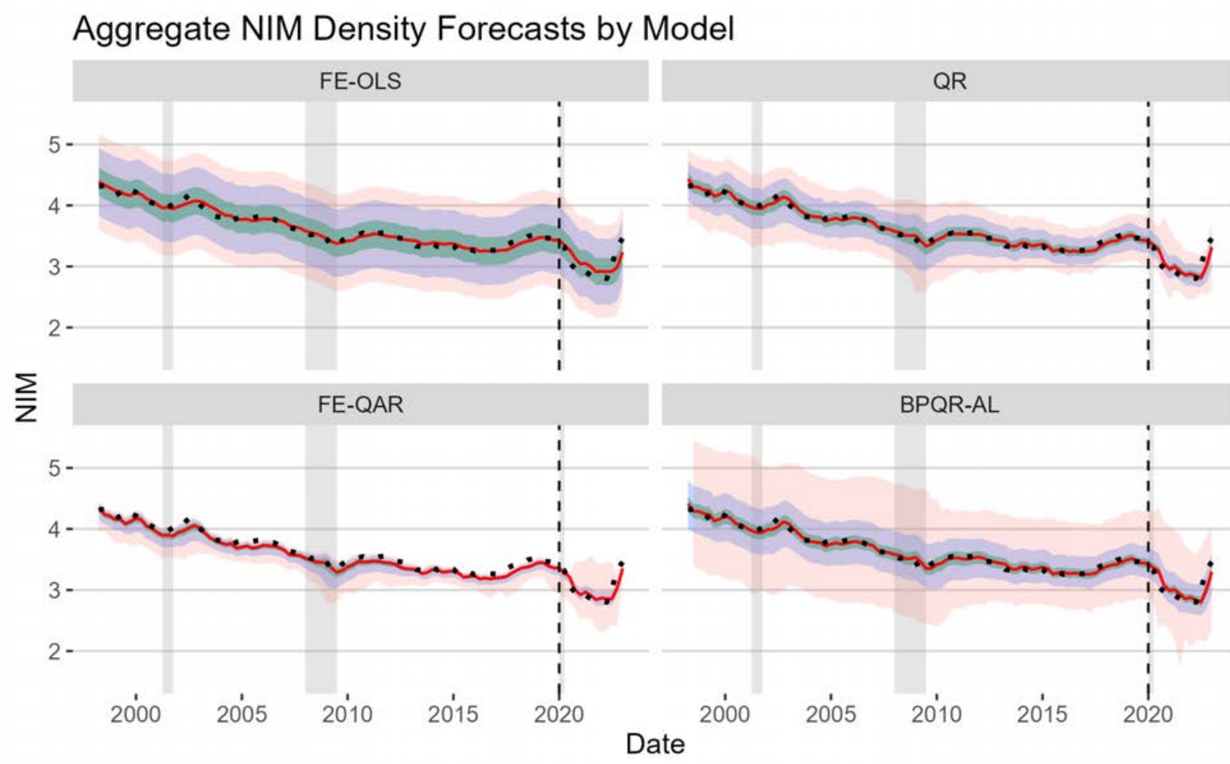

Figure 18: Aggregate NIM density forecasts by model. Plot includes fixed-effects OLS (FE-OLS), pooled quantile regression (QR), fixed-effects quantile autoregression (FE-QAR) and Bayesian panel quantile regression with adaptive LASSO (BPQR-AL). Density forecasts for the QR and BPQR-AL models are estimated from the model-specific quantile forecasts fitted to the skewed-t distribution proposed by Azzalini and Capitanio (2003). Density forecasts aggregated across banks in each period using Vincentization. Shaded areas represent 1st, 5th, 25th, 75th, 95th, and 99th percentiles predictions, solid red line represents median prediction, and dashed black line represents average actual NIM outcome. Dashed vertical line represents final in-sample prediction; one-quarter-ahead pseudo out-of-sample predictions begin in 2020 Q1. Vertical grey areas represent NBER-defined US recessions.

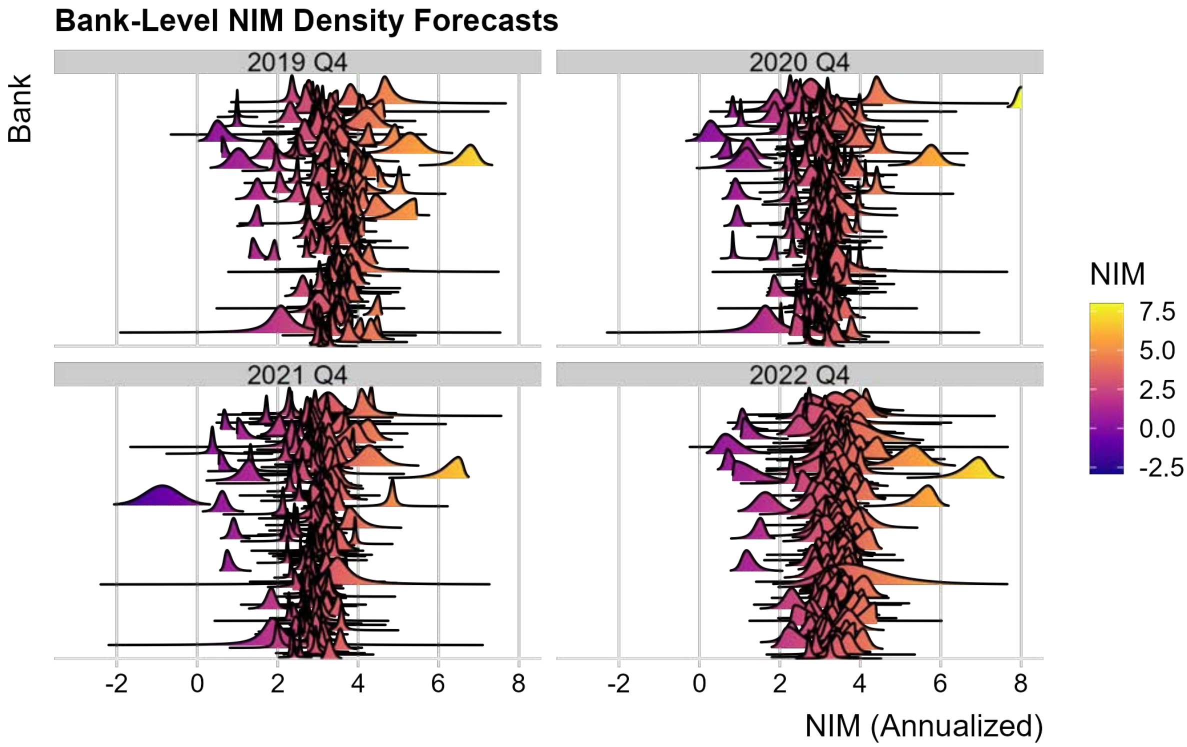

Figure 19: Individual bank density forecasts based on BPQR-AL pseudo out-of-sample estimations. Density forecasts are estimated from the BPQR-AL quantile forecasts fitted to the skewed-t distribution proposed by Azzalini and Capitanio (2003).

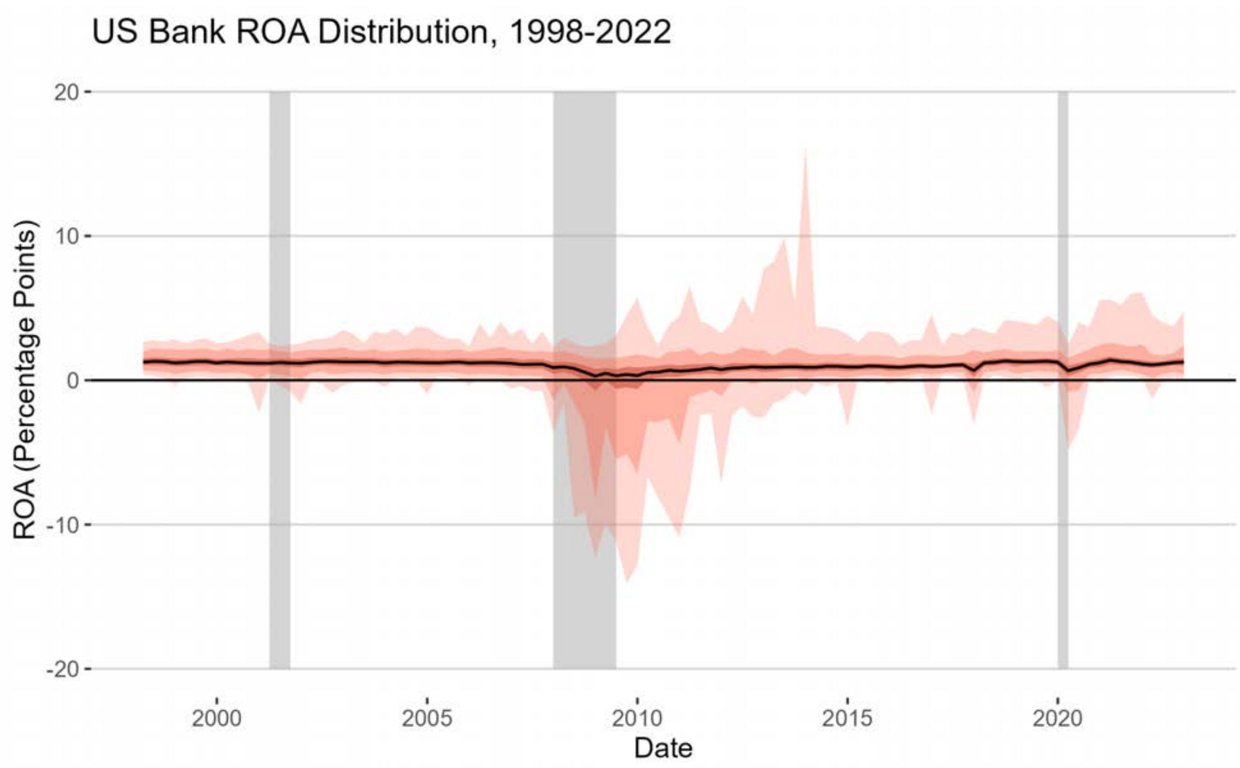

Figure 20: Return on assets (annualized) distribution of bank holding companies, 1998-2022. Shaded areas represent 1st, 5th, 25th, 75th, 95th, and 99th percentiles. Sample excludes bank holding companies with zero balances of total assets, interest-earning assets, or loans. Source: Merger-Adjusted FRY9-c Data (MAY9c)

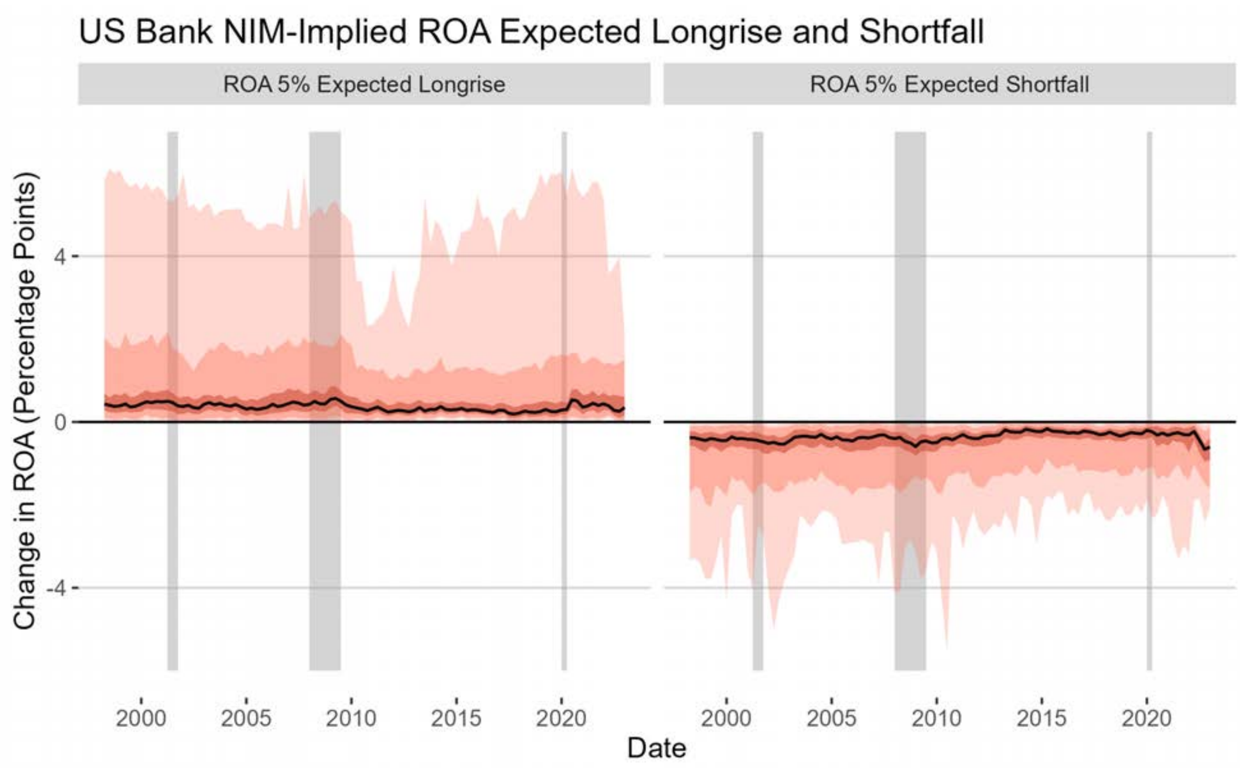

Figure 21: 5 percent expected longrise and shortfall of the change in bank return on assets (ROA) implied by individual bank BPQR-AL NIM forecasts. In-sample one-quarter-ahead predictions from 1998-2019, pseudo out-of-sample one-quarter-ahead predictions begin in 2020 Q1 and end in 2022 Q4. Shaded areas represent 1st, 5th, 25th, 75th, 95th, and 99th percentiles. Density forecasts are estimated from the BPQR-AL quantile forecasts fitted to the skewed-t distribution proposed by Azzalini and Capitanio (2003)

14 Appendix

14.1 BPQR-AL Trace and ACF Plots

To examine the convergence of the Gibbs sampling algorithm used to estimate the BPQR-AL model, I plot the trace plots and autocorrelation function plots for the parameters in the model. I produce a set of such plots for each of the five conditional quantiles estimated, being the 5th, 25th, 50th, 75th, and 95th quantiles. The trace plots show that the Gibbs sampler generally converges after the burn-in draws are discarded. The autocorrelation function plots show that the parameters of the model become reasonably stationary, although some autocorrelation remains for some of the balance sheet parameters.

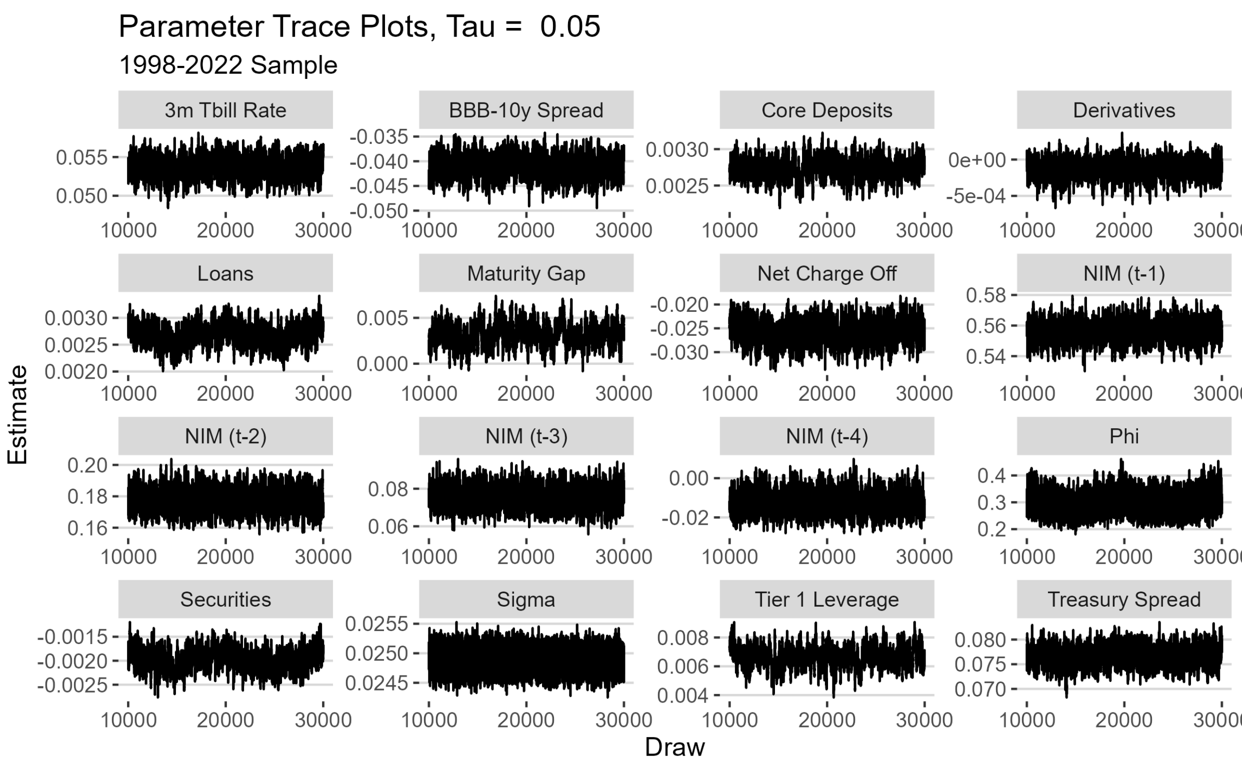

Figure 22: BPQR-AL Trace Plots from estimation on the full sample of banks spanning 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws.

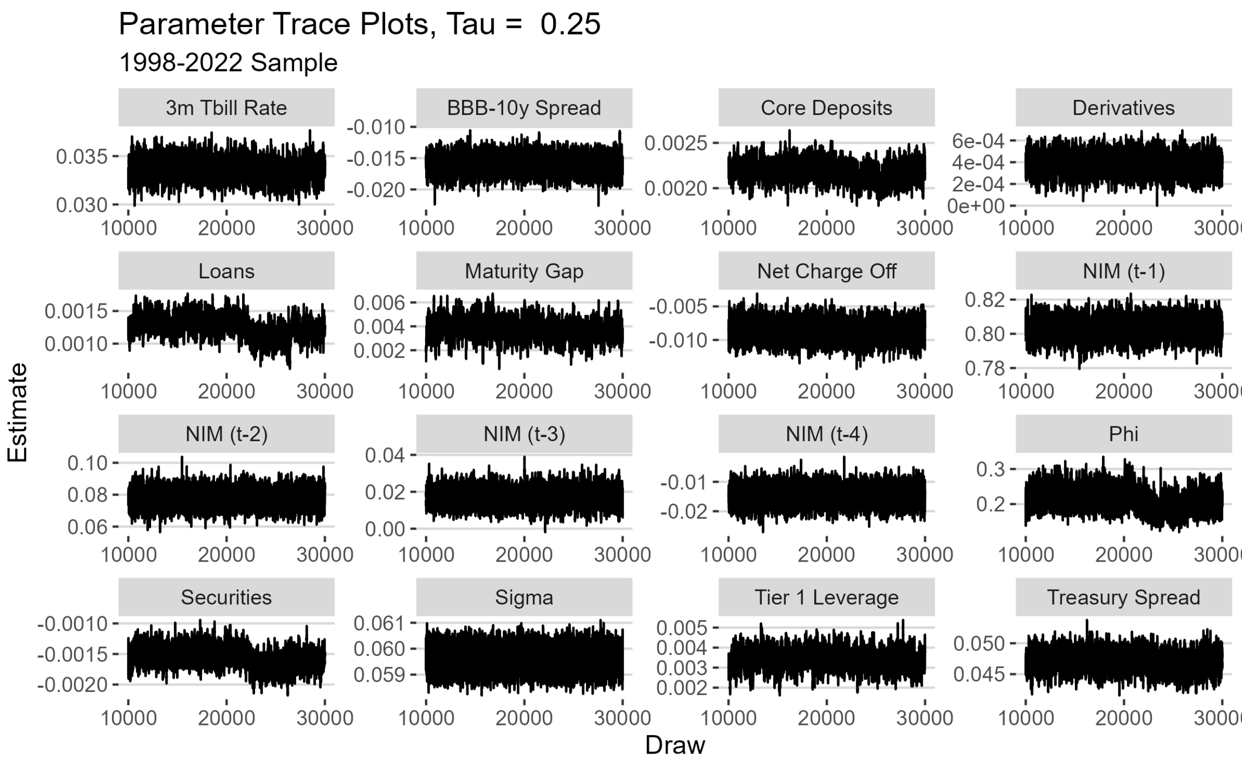

Figure 23: BPQR-AL Trace Plots from estimation on the full sample of banks spanning 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws.

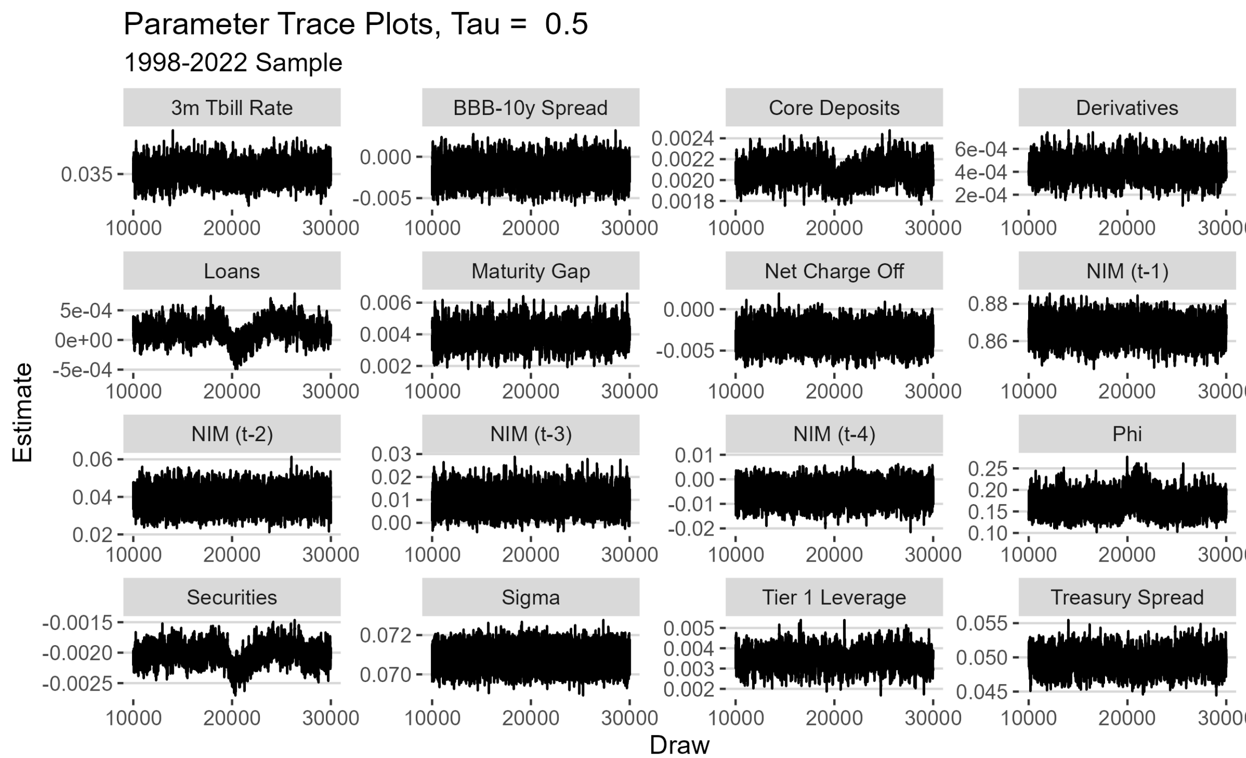

Figure 24: BPQR-AL Trace Plots from estimation on the full sample of banks spanning 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws.

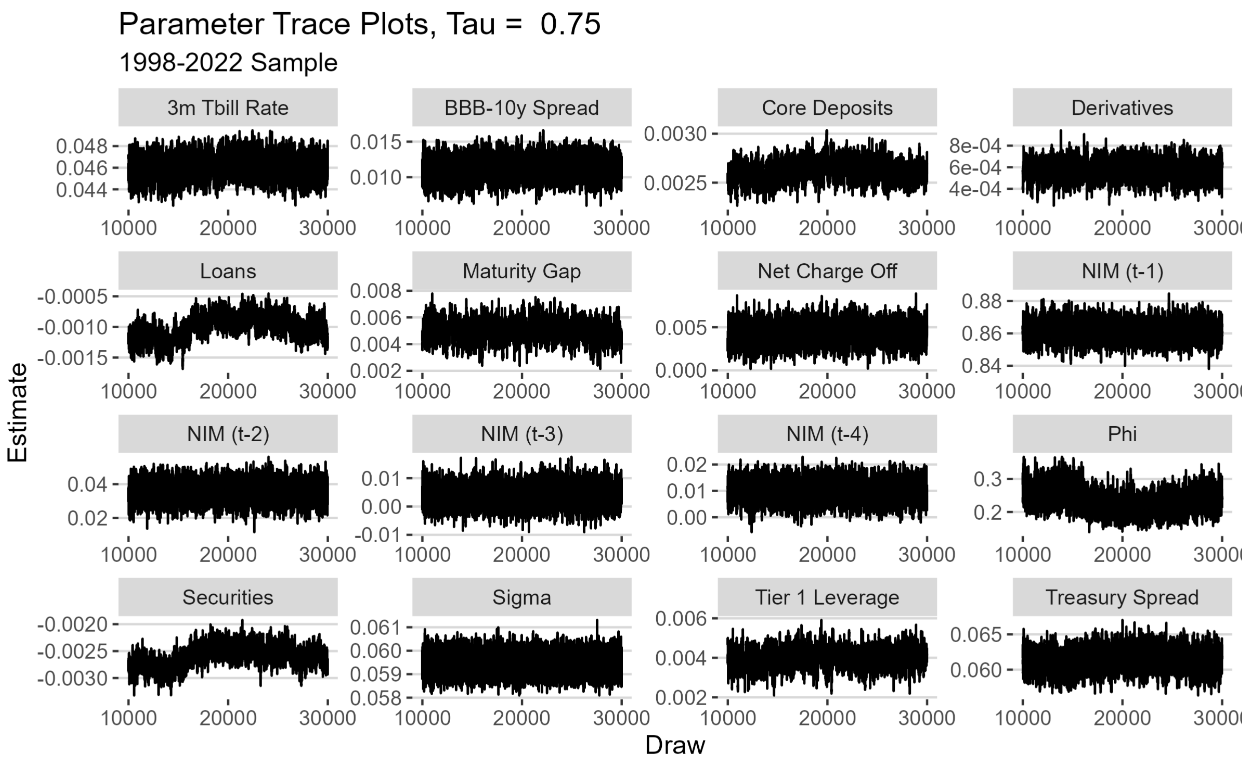

Figure 25: BPQR-AL Trace Plots from estimation on the full sample of banks spanning 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws.

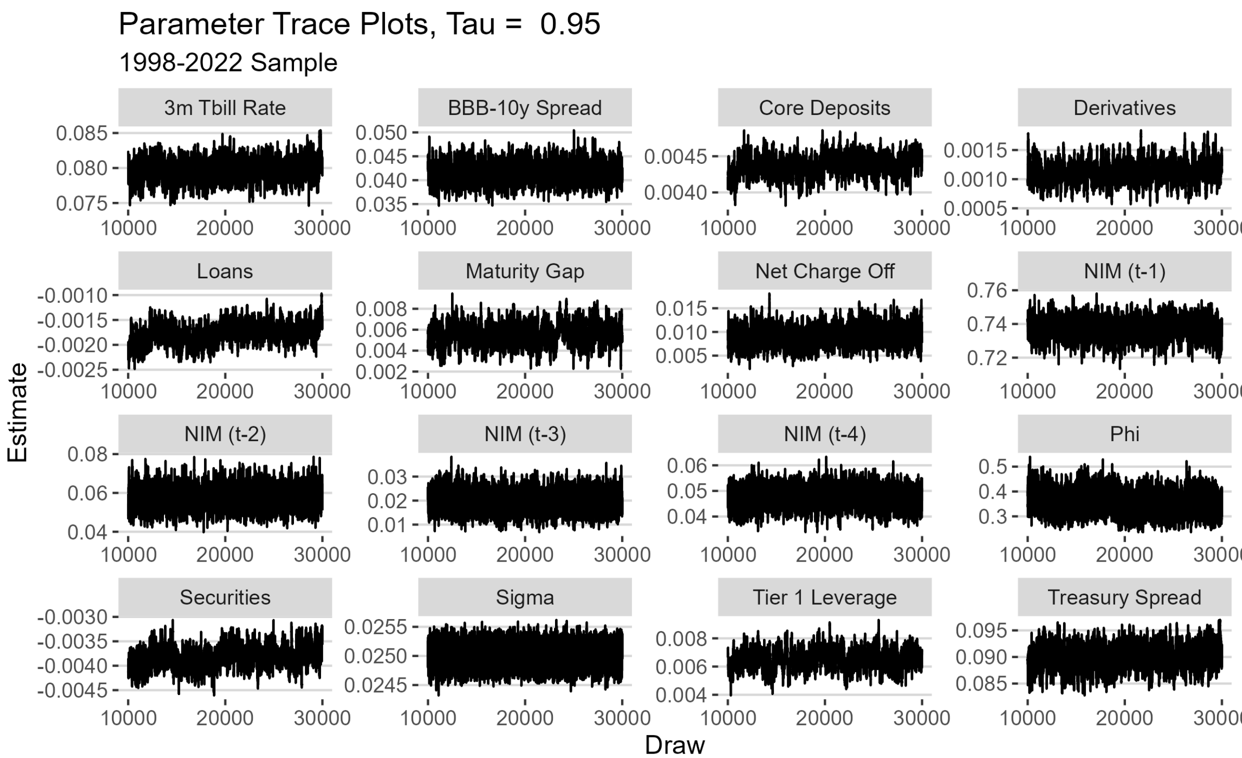

Figure 26: BPQR-AL Trace Plots from estimation on the full sample of banks spanning 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws.

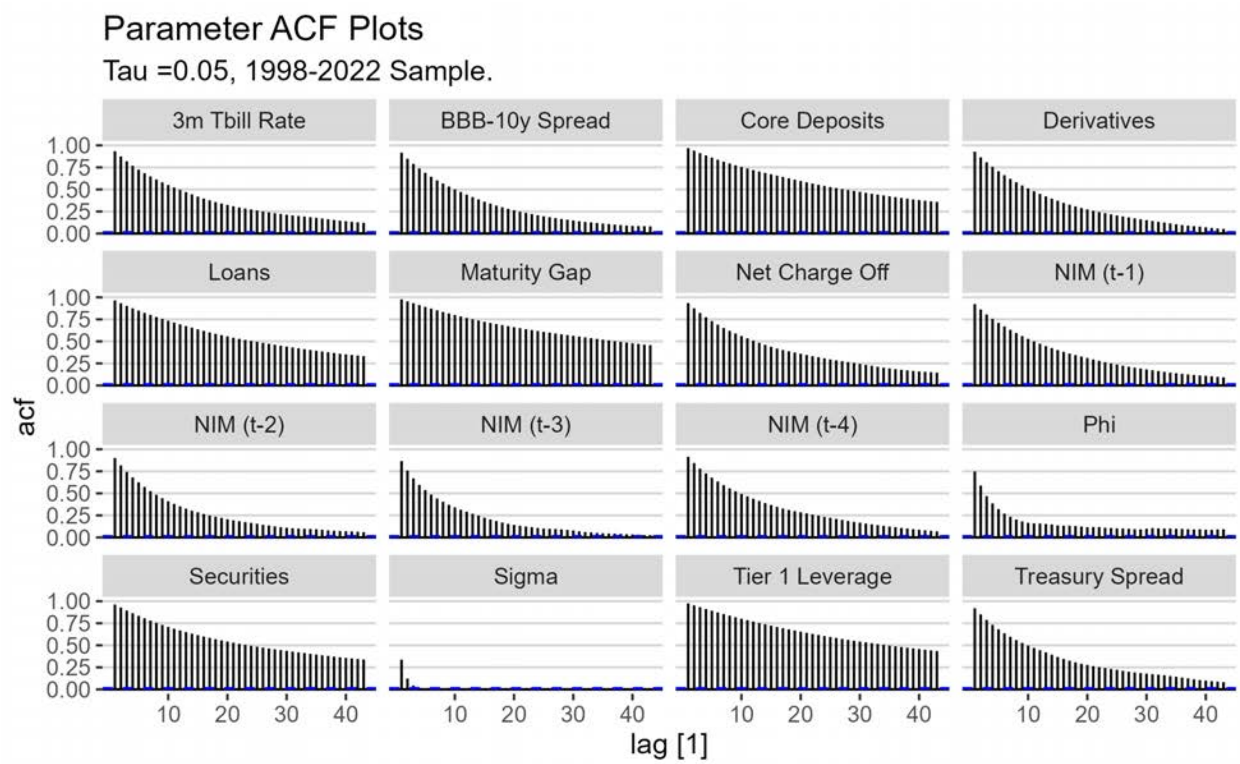

Figure 27: BPQR-AL ACF Plots from estimation on the full sample of banks spanning 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws.

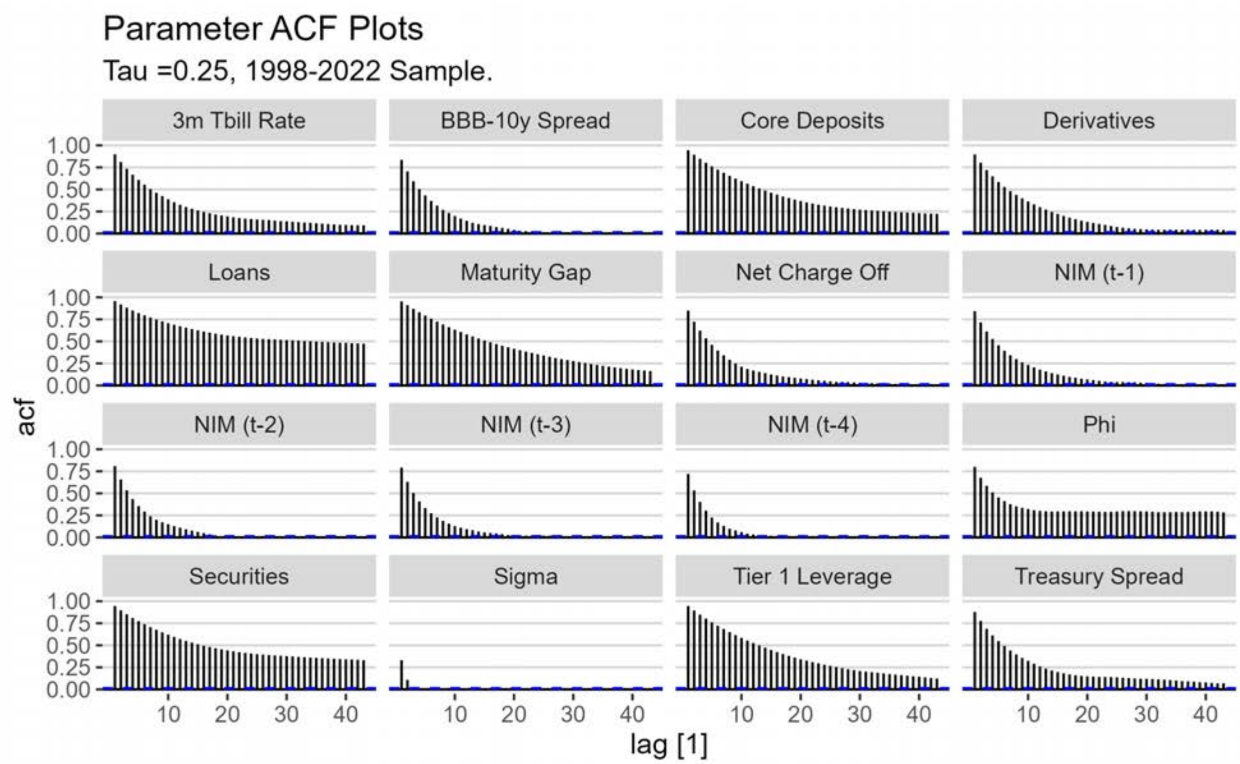

Figure 28: BPQR-AL ACF Plots from estimation on the full sample of banks spanning 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws.

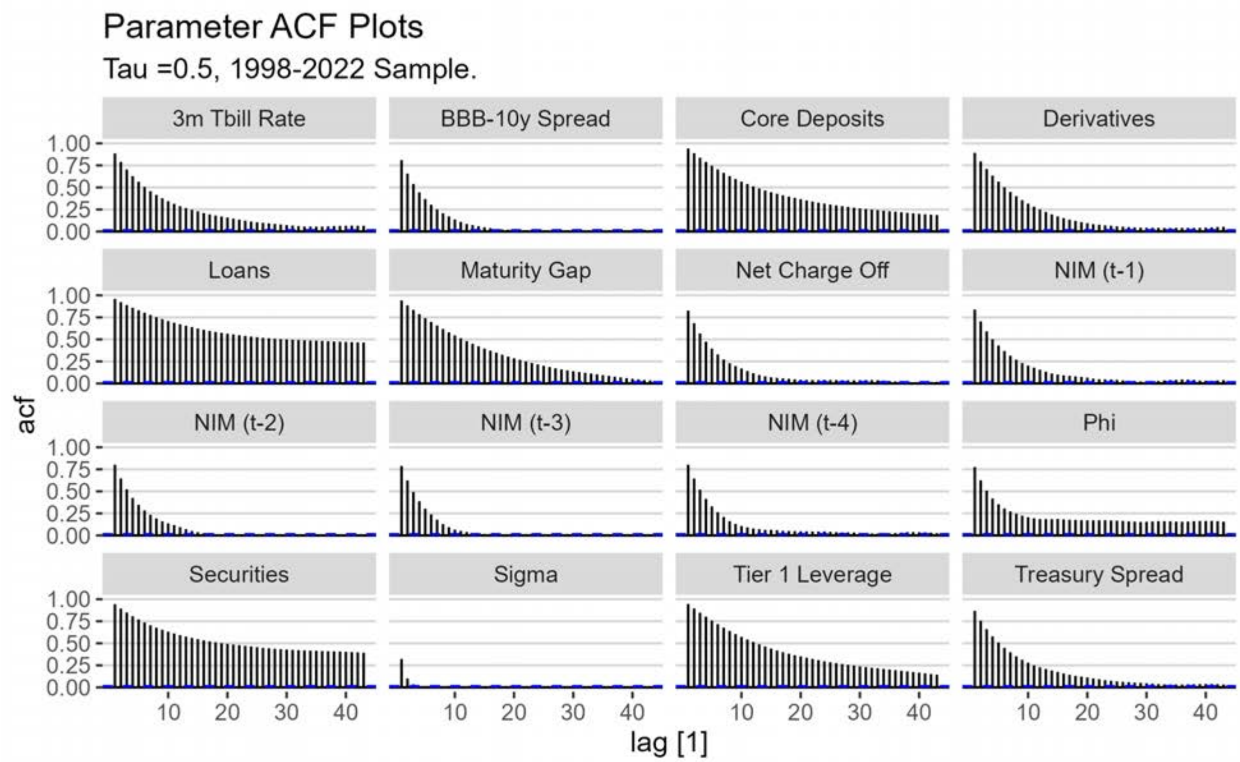

Figure 29: BPQR-AL ACF Plots from estimation on the full sample of banks spanning 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws.

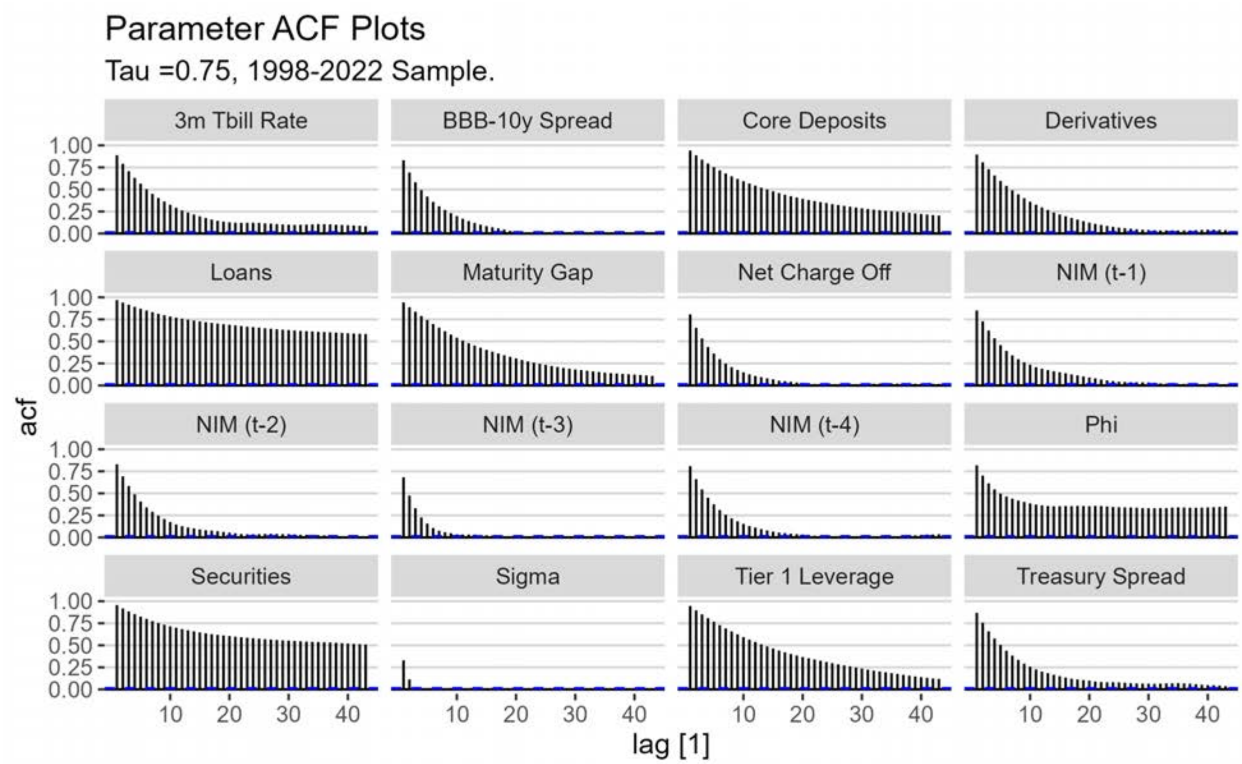

Figure 30: BPQR-AL ACF Plots from estimation on the full sample of banks spanning 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws.

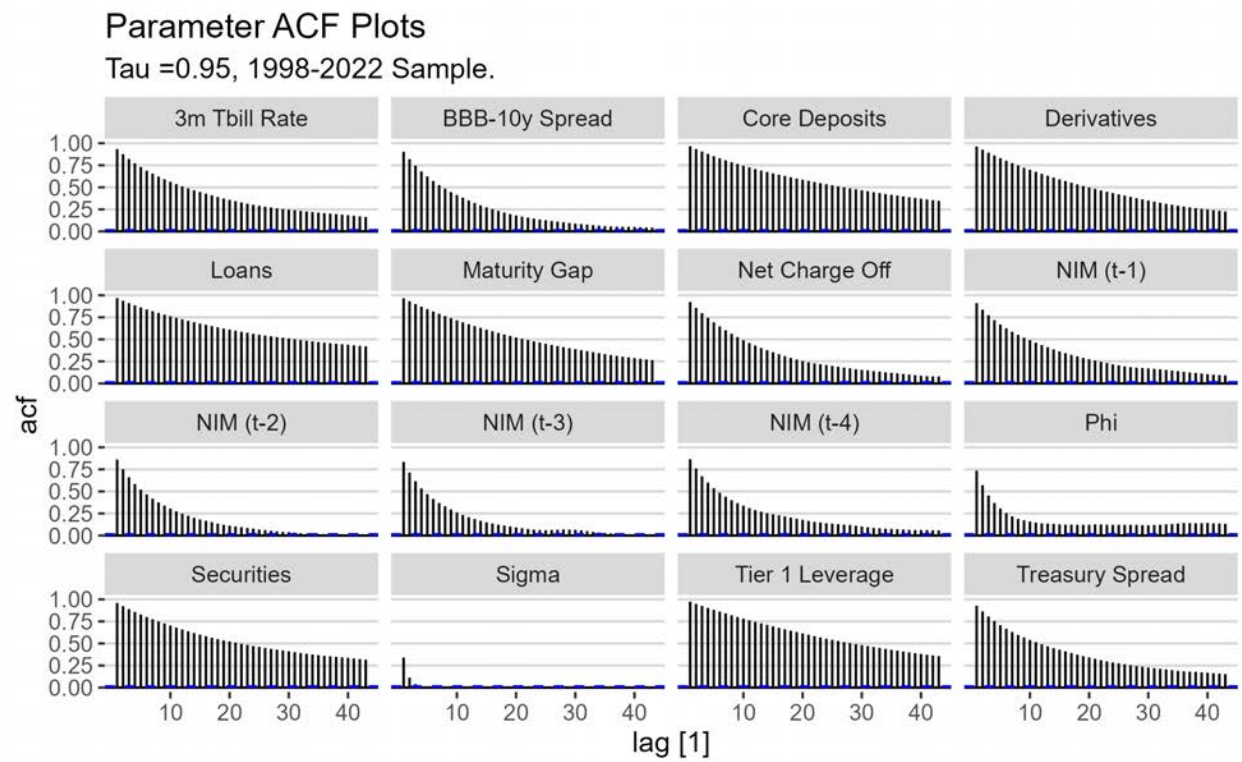

Figure 31: BPQR-AL ACF Plots from estimation on the full sample of banks spanning 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles. 30,000 total draws taken from MCMC Gibbs sampling algorithm with 10,000 burn-in draws.