

Three Essays in Empirical Banking and Finance

Nicholas Fritsch

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Department of Economics

The University of Strathclyde

Glasgow

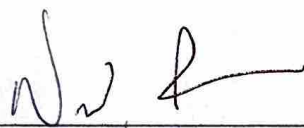
September 25, 2026

The views expressed in this thesis are those of the author and not necessarily those of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

Declaration of Authenticity and Author's Rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed: 

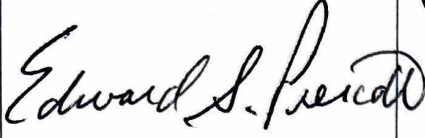
Date: 7/19/2025

Current Publication Status and Statement of Contribution of Co-Authors

Earlier versions of Chapters 3 and 4 have been released for working papers, Chapter 5 is planned to be released as a working paper after the viva. Chapter 3 is currently under editorial review. Chapter 4 is planned to be submitted for publication after the viva.

The undersigned hereby certify that:

1. they meet the criteria for co-authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the working paper in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the working paper;
3. there are no other authors of the publication according to these criteria; and
4. there are no potential conflicts of interest that require to be disclosed to (a) granting bodies, (b) the editor or publisher of any journal or other publication.

Contributor	Statement of Contribution	Publication title and date of publication or status
Edward S. Prescott 	Conceptualization, co-writing, methodology	<i>Macroeconomic Parameter Instability in Auto Loan Loss Models</i> , submitted to Federal Reserve Bank of Cleveland Working Paper Series, link not yet available.
Nicholas Fritsch 	Conceptualization, co-writing, methodology, data work, coding	

Contributor	Statement of Contribution	Publication title and date of publication or status
James Mitchell 	Conceptualization, co-writing, methodology	<i>Probabilistic Anchoring of Inflation Expectations</i>
Nicholas Fritsch 	Conceptualization, Co-writing, methodology, data work, coding	

Abstract

This thesis consists of three essays examining key issues in empirical macroeconomics and finance, with particular focus on interest rate transmission, financial stability, and inflation expectations.

The first essay investigates the relationship between net interest margins (NIM) and interest rates across the conditional distribution. Using Bayesian dynamic panel quantile regression, I document increased sensitivity to interest rates in the tails of the conditional NIM distribution during the post-2020 period, partly a result of bank interest expenses constrained by the zero lower bound. These findings have implications for measurement of interest rate risk and financial stability.

The second essay examines macroeconomic instability and credit risk transitions in auto loans using a discrete-time Markov transition framework. Using loan-level data, we document persistent declines in the sensitivity of default probabilities to unemployment shocks during both the Global Financial Crisis and Covid recessions. The findings have important implications for stress testing models and loss forecasting practices.

The third essay develops a probabilistic measure of inflation expectations anchoring that extends beyond testing whether mean expectations equal the target. Using risk-neutral densities from inflation options and Bayesian functional regression methods, we show that anchoring across the probabilistic distribution deteriorated in two waves with fundamentally different characteristics. We find that the 2024-2025 wave, reflecting elevated high-inflation tail probabilities, would be missed entirely by more traditional point-forecast methods.

Contents

Abstract	5
1 Acknowledgements	11
2 Introduction	13
2.1 Motivation	13
2.2 Overview of the Thesis	14
2.2.1 Chapter 1: Tail Sensitivity of US Bank Net Interest Margins	14
2.2.2 Chapter 2: Macroeconomic Instability and Credit Risk Transitions in Auto Loans	15
2.2.3 Chapter 3: Probabilistic Anchoring of Inflation Expectations	16
2.3 Contribution	18
2.4 Organization	19
3 Tail Sensitivity of US Bank Net Interest Margins	21
3.1 Introduction	21
3.2 Literature	24
3.3 Bank Net Interest Margins	26
3.4 Econometric Framework	29
3.5 Data	34
3.6 Estimation Results	35
3.6.1 Full Sample Results	35
3.6.2 Time Subsample Results	37
3.6.3 Income and Expense Decomposition	39
3.6.4 Results by Bank Size	41
3.7 Density Forecasting	42
3.7.1 Density Forecast Construction	43
3.7.2 Density Forecast Evaluation	44
3.8 Measures of Risk	47
3.9 Conclusion	49
3.10 Tables	52
3.11 Figures	56
3.12 Internet Appendix	71
3.12.1 Sample Attrition	71
3.12.2 Adaptive LASSO	71

3.12.3	NIM Numerator Decomposition	73
3.12.4	Loan and Security Concentrations in the Cross-Section	74
3.12.5	Bank Size Results	75
3.12.6	Robustness Checks	77
3.12.7	Alternative Model Forecast Performance Over Time	82
4	Macroeconomic Parameter Instability in Auto Loan Loss Models	83
4.1	Introduction	83
4.2	Background on the Auto Loan Market	85
4.3	Literature	86
4.4	Data	88
4.5	Empirical Framework	90
4.5.1	Multi-Period-Ahead Forecasts of Default	94
4.6	Model Estimation	94
4.6.1	Transitions from the Current State	95
4.6.2	Transitions from the Delinquent State	96
4.7	Rolling Regressions, Predicted Transition Probabilities, and Coefficient Instability	97
4.7.1	Forbearance	101
4.7.2	Controlling for 2020 Interventions	101
4.8	Cumulative Forecasts	102
4.8.1	Borrower Risk Score and Lender Channels	104
4.8.2	Oaxaca-Blinder Decomposition	105
4.8.3	Comparison to Binomial Discrete-Time Hazard Model	107
4.9	Conclusion	110
4.10	Tables	126
4.11	Figures	132
4.12	Appendix	156
4.12.1	Hazard Specification	156
4.12.2	Regression Results: 2000-2006	157
4.12.3	Forecasts and Parameter Instability by Lender Type	158
4.12.4	Forecasts and Parameter Instability with Alternative Macroeconomic Specifications	159
4.13	Appendix Tables and Figures	160
5	Probabilistic Anchoring of Inflation Expectations	173
5.1	Introduction	173
5.2	Data	179
5.2.1	Option-Implied Risk-Neutral Inflation Densities	179
5.2.2	Macroeconomic surprises	183
5.3	Probabilistic Anchoring of Inflation Expectations	184
5.3.1	Two Dimensions of Probabilistic Anchoring	188
5.3.2	Regional Anchoring: A Policy-Relevant Relaxation	189
5.4	Econometric Framework	191
5.4.1	Bayes Space and Centered Log Ratio Transformation	192

5.4.2	Testing Probabilistic Anchoring via Functional Regression	195
5.4.3	Functional Principal Components Analysis	199
5.5	Empirical Results: Three Findings About Probabilistic Anchoring	203
5.5.1	Fact 1: Total Anchoring Deteriorated in Two Waves	203
5.5.2	Fact 2: The Nature of Unanchoring Changed from Good to Bad	205
5.5.3	Labor Market Effects on Upper-Tail Unanchoring	211
5.5.4	Comparison to Traditional Metrics	216
5.6	Conclusions	218
	Online Appendix	220
5.7	Additional Figures	222
6	Conclusion	231
6.1	Summary of Findings	231
6.2	Common Themes	232
6.2.1	Empirical Instability	232
6.2.2	The Importance of Distributional Analysis	233
6.3	Policy Implications	234
6.4	Concluding Remarks	235

Chapter 1

Acknowledgements

I am deeply grateful to the many individuals who have contributed to this dissertation and my development as a researcher.

First and foremost, I would like to thank my advisors, Gary Koop and Stuart McIntyre, for their invaluable guidance, support, and mentorship throughout this process. Their insights and feedback have been instrumental in shaping this work. I owe much of my development as a researcher over the duration of the PhD program to them.

I am also indebted to Edward S. Prescott and James Mitchell for their collaboration and contributions to the coauthored chapters in this dissertation. Working with them has been a privilege and an education, and I have gained a wealth of research experience from our work together.

I am very grateful to Ben Craig for his additional guidance and support, which further refined my research and expanded my understanding of the field. Despite his status as emeritus faculty at the Federal Reserve Bank of Cleveland, Ben was generous enough to meet with me regularly for the duration of my PhD and offered his wealth of technical and practical expertise as I progressed through my chapters.

I would also like to thank the many economists with whom I have had useful discussions over the course of this research. These conversations have challenged my thinking, broadened

my perspective, and contributed significantly to the development of the ideas presented here. The Research staff at the Federal Reserve Bank of Cleveland have been especially helpful, particularly Paola Boel, Daniel Carroll, Paolo Gelain, Baris Kaymak, Lara Lowenstein, Rob Rich, Jan-Peter Siedlarek, Randall Verbrugge, Saeed Zaman, and Pete Zimmerman. Faculty in the Economics Department of the University of Strathclyde have been extremely helpful with comments and suggestions during my PhD journey, particularly Niko Hauzenberger, Ping Wu, Sharada Davidson and Julia Darby. Given that I completed this PhD while working at the Federal Reserve Bank of Cleveland, I also thank my former supervisor Denise Duffy for supporting me in pursuing the PhD.

Finally, any errors or omissions in this dissertation are entirely my own.

Chapter 2

Introduction

2.1 Motivation

The period from 2000 to 2025 has been marked by macroeconomic volatility and policy interventions that have challenged assumptions around key relationships in financial markets and the macroeconomy. The Global Financial Crisis of 2007-2009, the Covid-19 pandemic of 2020-2021, and the subsequent inflation surge of 2022-2024 have each tested the resilience of financial institutions and the stability of financial forecasting models. A common thread running through these essays is the question of structural stability—whether relationships that appear stable during normal times remain reliable during periods of extreme stress and policy intervention in aggregate or across the empirical distribution.

In the first chapter, understanding interest rate risk in bank profitability depends on stable relationships between interest rates and interest-sensitive income and expenses. The second chapter addresses stress testing models used by banking regulators that rely on stable relationships between macroeconomic variables and loan performance. The final chapter explores distributional anchoring of inflation expectations, relaxing the common focus on point estimates. If these relationships prove unstable in aggregate or across the empirical distribution, then models and policy frameworks built upon them may provide misleading

results.

The three essays in this thesis contribute to our understanding of these issues by applying novel empirical methods to datasets spanning multiple business cycles and documenting how traditional relationships have evolved over the past quarter century.

2.2 Overview of the Thesis

2.2.1 Chapter 1: Tail Sensitivity of US Bank Net Interest Margins

The first essay examines the relationship between banks' net interest margins (NIM) and interest rates across the conditional distribution. Net interest margin—the difference between interest income earned on assets and interest expense paid on liabilities, expressed as a percentage of earning assets—is a primary determinant of bank profitability. Understanding how interest rate movements affect aggregate and individual bank margins is crucial for assessing the financial stability implications of policy changes.

This relationship has received renewed attention following the rapid decline in short-term interest rates during the pandemic and the subsequent sharp increase during 2022-2023. The conventional wisdom holds that, for example, falling rates can be costly for banks as assets originate or reprice at relatively lower rates, leading to reduced cash flows. The academic literature has documented that banks can offset this by decreasing deposit costs along with market rates to maintain a comparable spread between asset and deposit interest rates (Drechsler et al. (2021), Williams (2020)) However, the extent to which banks are able to successfully decrease deposit costs depend on how close to the zero lower bound their current deposit costs are. When banks lose the ability to reprice deposits further downward, nonlinearities can occur in declines to net interest margins. Standard linear approaches to estimate the impact of interest rate changes on margins in such cases can mask distributional heterogeneity.

The chapter examines these issues empirically using bank-level panel data, documenting

how net interest margins respond to interest rate changes and exploring heterogeneity across banks and over time. Using the novel Bayesian panel quantile regression developed by Aghamohammadi and Mohammadi (2017), I examine how changes in interest rates have heterogeneous impacts on banks across the conditional distribution. The model deals with the incidental parameter problem, allowing for estimation of both individual fixed effects and covariates that vary across the conditional distribution. This allows me to highlight how interest rate effects across the conditional margin distribution vary across banks and over time.

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, and lower NIM tail outcomes that are partly a result of bank interest expenses constrained by the zero lower bound. These effects differ across bank sizes, where larger banks explain more of the increased sensitivity to interest rates in the adverse tails of the distribution. Density forecast evaluation shows that the model forecasts outperform frequentist linear benchmark models, suggesting that modeling the conditional distribution may become more useful after periods of interest rate volatility. These findings have implications for understanding the differential impact of monetary policy across the banking system and for assessing financial stability risks.

2.2.2 Chapter 2: Macroeconomic Instability and Credit Risk Transitions in Auto Loans

The second essay, coauthored with Edward S. Prescott, examines the stability of credit risk models in the context of auto loans, the second-largest consumer credit market in the United States with \$1.56 trillion outstanding as of 2025. Using loan-level data from the FRBNY Consumer Credit Panel covering 2000-2025, we estimate a discrete-time Markov transition model that jointly accounts for movements through delinquency, default, prepayment, and cure states. This approach represents a methodological contribution to the auto lending

literature, which has typically focused on terminal default or prepayment probabilities while ignoring the probabilistic dynamics of transitions across financial states of loan performance (for example, Agarwal (2008); Heitfield and Sabarwal (2004); Wu et al. (2018)).

We document that the sensitivity of default probabilities to unemployment shocks—a key parameter in stress testing and loan loss forecasting models—has declined persistently following both the Global Financial Crisis and the Covid recession. In 2006, a one percentage point increase in unemployment was associated with a 16% increase in default risk. By post-2020, this effect had fallen to just 3%. Using rolling pseudo out-of-sample forecasts, we show that models estimated before the pandemic would have overstated two-year cumulative defaults by 100 basis points (25%) during 2020-2021.

Importantly, this instability persists even after controlling for accounts observed to go into forbearance and also pandemic-period dummy variables that correlate with other government transfers. Using an Oaxaca-Blinder decomposition, we demonstrate that the forecast errors are driven primarily by changes in the macroeconomic relationships themselves rather than changes in borrower composition. The essay also documents significant heterogeneity across lender types, with loans originated by auto finance companies, banks, and credit unions exhibiting different risk characteristics and different sensitivities to macroeconomic shocks.

These findings have direct implications for the stress testing models used by banking regulators as well as for the internal loss forecasting models used by auto lenders. The persistent decline in unemployment sensitivity suggests that these models may systematically overstate credit losses during future recessions, potentially leading to inefficiently high capital requirements or overly conservative lending standards.

2.2.3 Chapter 3: Probabilistic Anchoring of Inflation Expectations

The third essay, co-authored with James Mitchell, examines anchoring of inflation expectations. Existing research typically focus on point expectations—testing whether the mean or median forecast equals the 2% target or whether it responds to macroeconomic news. This

approach misses a crucial dimension: the full probability distribution of inflation expectations. Our framework generalizes an approach in the literature that measures deviation of inflation expectations from a target point or region (for example, Kumar et al., 2015, Grishchenko et al., 2019, Kitsul and Wright, 2013). Unlike these papers, we examine deviations pointwise across the distribution, as well as deviations that are aggregated across regions of the distribution.

We develop a probabilistic measure of inflation expectations anchoring that extends traditional tests to the entire density. Using risk-neutral densities extracted from inflation options over 2012-2025, we apply Bayesian function-on-scalar regression in Bayes space with centered log-ratio transformations to test whether densities are both unconditionally equal to the target density and conditionally orthogonal to macroeconomic news. This functional data analysis approach allows us to identify precisely where in the distribution—center versus tails—macroeconomic surprises have their effects. Our analysis examines deviations from a policy maker’s “target density”—the complete probability distribution that reflects not just the central bank’s inflation target but also the tolerance for deviations around it. Central banks universally recognize that inflation will not remain exactly at target; the FOMC aims for inflation to “average 2 percent over time,” the Bank of Canada specifies a “control range” of 1-3%, and the ECB targets inflation “close to 2 percent.” These statements implicitly define a target density, not merely a target point.

Given we focus on the full distribution, we define regions of the densities such that significant movements over time could represent “good” or “bad” unanchoring, the distinction lying in where along the support the probability mass is shifting. Our findings reveal that anchoring deteriorated in two waves with fundamentally different characteristics. The 2018-2020 wave reflected “good” regional anchoring, with excess probability mass concentrated very close to target. In contrast, the 2024-2025 wave reflects “bad” upper-tail unanchoring, with elevated probabilities of high inflation outcomes despite stable mean expectations. This upper-tail unanchoring is correlated with labor market surprises and would be completely

missed by traditional point-forecast metrics that focus only on the mean.

These findings have important implications for how central banks should monitor and communicate about inflation expectations. If markets assign significant probability to inflation running persistently above target—even when mean expectations remain near 2%—then expectations are not truly anchored in the sense that matters for welfare and policy. In such cases, central banks should expand their monitoring frameworks to include distributional measures of anchoring.

2.3 Contribution

These three essays make several contributions to the empirical banking and finance literature. Methodologically, they demonstrate the value of moving beyond point estimates to examine full distributions (Chapter 1 and 3) and full state-space dynamics (Chapter 2). They highlight the importance of examining heterogeneity across institutions (Chapter 1 and 2) rather than assuming representative agents or representative banks. And they show how microdata—whether bank-level panel data, consumer loan-level observations, or densities implied by traders in derivative markets—can reveal heterogeneity and dynamics that are obscured in aggregate relationships.

The essays document significant instability and heterogeneity in relationships that are central to policy models and risk management practices. The heterogeneous response of bank margins to interest rate changes (Chapter 1) has implications for understanding interest rate risk for bank earnings and financial stability. The declining sensitivity of auto loan defaults to unemployment (Chapter 2) challenges the assumptions underlying stress testing frameworks. And the evidence of distributional unanchoring of inflation expectations (Chapter 3) suggests that point-forecast metrics provide an incomplete picture of expectation formation.

From a policy perspective, these findings suggest caution in relying heavily on reduced form models estimated during historical periods. They also suggest that policymakers and

risk managers should place greater weight on examining full distributions rather than conditional means, should account for heterogeneity across institutions and borrowers, and should regularly test for structural breaks and parameter instability. More broadly, they highlight the challenges of conducting empirical finance research and policy analysis in an environment where the very relationships we seek to understand appear to be evolving over time.

2.4 Organization

The remainder of this thesis is organized as follows. Chapter 1 examines net interest margins and monetary policy. Chapter 2 presents the analysis of auto loan credit risk transitions. Chapter 3 develops and applies the probabilistic measure of inflation expectations anchoring. Chapter 4 concludes with a discussion of common themes, policy implications, and directions for future research.

Chapter 3

Tail Sensitivity of US Bank Net Interest Margins

3.1 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

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

a disproportionately large amount of total assets.

There exist differing views in the academic literature on the ability of banks to hedge macroeconomic interest rate impacts on bank NIMs. Past work has also 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 meaningful differences in sensitivity to interest rates, credit risk, and persistence across the conditional NIM distribution, which of these factors helps to explain the recent break in NIM stability, and does estimation of heterogeneous sensitivities improve predicted outcomes? It is 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 using a Bayesian dynamic panel quantile regression framework, and evaluating the resulting density forecasts against benchmark models using proper scoring rules of the density predictions. Covas et al. (2014) find that a dynamic panel quantile regression framework estimated on the largest banks subject to US Dodd-Frank Act stress testing requirements 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. The features of such a model may be useful when estimating NIM sensitivities to interest rates and predicting NIM outcomes when interest rates are volatile. Unlike Covas et al. (2014), 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.

I estimate the model on quarterly regulatory filing data from a sample of 313 US bank holding companies (BHC) during 1998-2022. 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. Additional analysis finds that banks' NIM sensitivity to interest rates has increased since 2020 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. These increased tail sensitivities are economically meaningful for banks when interest rate or credit risk movements are large. I also find evidence that interest expenses for banks in the lower tails of the conditional NIM distribution were constrained when short-term interest rates were near the zero lower bound during 2020-2021, exacerbating lower tail outcomes.

I use the conditional quantile estimates to produce full density forecasts for each bank in the sample. I find that the Bayesian dynamic panel quantile regression framework statistically outperforms fixed-effect OLS, pooled quantile regression, and the fixed-effect quantile autoregressive model framework of Covas et al. (2014) in forecasting pseudo out-of-sample one-period ahead outcomes over the post-pandemic period. 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, which subsequently improved by the end of 2022. The Bayesian framework that I use also captures relatively larger upper and lower tail risks across aggregated NIM density forecasts relative to competing model frameworks while still placing the highest probabilities closest to the aggregated actual NIM outcomes. Expected shortfall estimations across the individual bank NIM density forecasts show that the size of the 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.

3.2 Literature

The first strand of relevant literature centers around 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 that summarize the yield curve, namely, factors that capture the level and slope (Abdymomunov et al. (2023); Alessandri and Nelson (2015); Borio et al. (2017); Covas et al. (2014); English et al. (2018); Hirtle et al. (2016)). The level of the yield curve, typically measured by short-term risk-free interest rates, captures bank profits from a larger increase in interest income relative to interest expense if banks are able to price deposits below market rates (Alessandri and Nelson (2015); Borio et al. (2017)). 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 et al. (2017), English et al. (2018), Samuelson (1945)). Theoretical and empirical work shows that an additional concern is the impact of credit spreads, because increased credit risk should create incentives for banks to increase loan risk pricing to offset potential losses in the long run, while banks shifting their lending toward less risky and lower-yielding assets in the short-run (Angbazo (1997), Hanweck and Ryu (2005), Zarruk and Madura (1992)). The academic literature has also documented the importance of own persistence of bank net interest margins, 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 (English (2002), Flannery (1981)).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 maturities in their asset and liability portfolios. English (2002) concludes that bank selection

²The importance of bank NIM persistence is also referenced by the Federal Reserve for its stress-testing models of bank revenues which are specified using four lags of the dependent variable.

of assets, liabilities, deposit rates, and hedging activities explains in part the stability of observable NIMs as part of an international study. Drechsler et al. (2021) examine banks' ability to hedge interest rate risk arising from maturity transformation as a result of market power over the deposit franchise, leading bank NIMs to be relatively stable during periods of changing interest rates. Williams (2020) comes to conclusions similar to those of as Drechsler et al. (2021) for the average bank but emphasizes that there is heterogeneity in the ability to hedge interest rate risk in the cross section. 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 et al. (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.

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 et al. (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. 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 et al. (2019) in constructing continuous density forecasts by fitting a parsimonious set of quantile forecasts to a skewed t -distribution. Unlike these studies, I focus exclusively on bank net interest margins for a large merger-adjusted sample of 313 US bank holding companies over a relatively long time period spanning 1998-2022, and I use a Bayesian approach

to panel data quantile regression to examine heterogeneity in the conditional distribution.

3.3 Bank Net Interest Margins

NIM is a measure of bank profits that are a function of prevailing interest rates, capturing profits from the interest rate spread earned on the maturity transformation role that banks play. NIM is defined as the difference between the interest income and interest expense generated by an institution, normalized by its interest-earning assets. For institution i at time t , NIM is computed as

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

NIM is an important source of revenue for banks, often representing a large share of gross income. For example, the interest income component of NIM accounts for roughly 80 percent of the total income, on average, for US bank holding companies since 1997 (henceforth “banks”).

The stability of bank NIM is apparent when compared to fluctuations in macroeconomic interest rates and credit risk. The left panel of Figure 3.1 shows average NIM for banks plotted against the 3-month Treasury bill and 10-year Treasury rates from 1998-2022. A long-term downward trend in NIM is visible over the period shown, explained in part by an overall decrease in the level of interest rates (Di Lucido et al. (2017)). Notably, 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 right panel 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 3.2 shows the cross-sectional distribution of bank NIM in terms of the median, interquartile range, and 5th-95th percentiles. There is evidence of some skewness in lower adverse tails of the distribution. Lower tail skewness is pronounced beginning in 2019 Q2, after which the distance between the median and lower 5th percentile reached a maximum of 89 basis points greater than the distance between the median and the upper 5th percentile by 2022 Q2. 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.

Decomposing bank NIM into interest income and interest expense shares of interest-earning assets, Figure 3.3 plots the distributions for US banks. The figure illustrates, importantly, that the ability of banks to adjust interest expense downward is constrained in periods when interest expenses are already near the zero-lower bound, and this appears to correlate with periods where the lower tail of NIM becomes more pronounced. 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 effect is even more pronounced 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 while interest expenses remained steady at nearly zero levels, having not increased enough prior to the pandemic period to allow for further decreases.

Turning to the denominator of NIM, interest-earning assets represents roughly 90 percent of total bank assets on average from 1998-2022.³ Figure 3.4 decomposes aggregate interest-earning assets into these subcomponents as a share of aggregate interest-earning assets from

³I compute interest-earning assets as the sum of reported quarterly averages of bank securities, loans, repos, and “other” interest-earning balances.

2001-2022.⁴ Loans typically represent the largest share of banks' interest-earning assets, followed by the securities portfolio. In the period following the pandemic, banks' loan balances declined while strong deposit growth was invested into investment securities and cash reserve balances, which offer relatively low rates of return.⁵ 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 (Ennis and Jarque (2021)). The increase in C&I lending also offered relatively low 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.⁶ While these portfolio changes occurred in aggregate, there exists large amount of cross-sectional heterogeneity in bank loan and security portfolio concentrations over time, representing a wide variety of business models and resulting sensitivities to interest rate changes.⁷

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

⁴In the FR Y-9c reporting forms, the "other" category of interest-earning assets is not reported until 2001.

⁵As 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>.

⁶PPP 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>.

⁷For illustrative evidence of this, see Appendix Figure 3.17

heterogeneity more pronounced in the post-pandemic period?

The theoretical and empirical banking literature suggests additional bank characteristics that can affect NIM. In the micro literature, NIM of a profit-maximizing bank may be a function of the bank’s balance sheet exposure to credit and interest rate risk (for example, Ho and Saunders (1981), Angbazo (1997)). Macrofinance literature notes maturity mismatch, reliance on relatively cheap “core” deposits, and the use of interest rate derivatives as bank-specific determinants of NIM (Drechsler et al. (2021), English et al. (2018)). I include a set of controls to capture relevant observable characteristics for each bank including Tier 1 leverage ratio, log of real total assets, loan and security shares of interest-earning assets, 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 et al. (2018).⁸ Table 3.1 presents summary statistics for all of the variables used in the empirical analysis.

3.4 Econometric Framework

As a starting point, we can specify a linear model of bank NIM in a dynamic panel fixed-effects OLS specification. Recent publications of the Federal Reserve Supervisory Stress Test methodology suggest that the FE-OLS model is in line with the general approach taken by regulators when modeling components of bank revenues, and thus serves as a natural benchmark.⁹ We can specify a linear dynamic panel model of bank NIM as

⁸See the appendix of English et al. (2018) that details construction of the maturity gap, <https://ars.els-cdn.com/content/image/1-s2.0-S0304393218302101-mm1.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 et al. (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.

⁹For example, see the 2024 Supervisory Stress Test Methodology published by the Federal Reserve Board of Governors, <https://www.federalreserve.gov/publications/files/2024-march-supervisory-stress-test-methodology.pdf>.

$$y_{i,t} = \alpha_i + \sum_{s=1}^k \zeta_s y_{i,t-s} + \mathbf{X}'_{i,t-1} \boldsymbol{\beta} + \mathbf{Z}'_t \boldsymbol{\gamma} + \epsilon_{i,t}, \quad i = 1, \dots, n, \quad t = 1, \dots, T, \quad N = n \times T, \quad (3.1)$$

where $y_{i,t}$ 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, $\mathbf{X}_{i,t-1}$ is a vector of balance sheet controls for time-varying heterogeneity across banks over time, and $\mathbf{Z}_t$ 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. Importantly, I treat interest rates and credit risk as exogenous outcomes and the focus is on relationships between observed macroeconomic fluctuations and bank NIMs.

Quantile regression provides a natural framework for examining conditional distributional sensitivities and allows for heterogeneous impacts of covariates across the conditional distribution, and can be used to recover the entire conditional distribution. The conditional quantile function of $y_{i,t}$ in Equation 3.1 is

$$Q_{y_{i,t}}(\tau_k | y_{i,t-1}, \dots, y_{i,t-k}, \mathbf{X}'_{i,t-1}, \mathbf{Z}'_t) = \alpha_i(\tau_k) + \sum_{s=1}^k \zeta_s(\tau_k) y_{i,t-s} + \mathbf{x}'_{i,t-1} \boldsymbol{\beta}(\tau_k) + \mathbf{Z}'_t \boldsymbol{\gamma}(\tau_k) \quad (3.2)$$

where τ_k is a quantile in $(0, 1)$.

Estimation a large number 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 well-known incidental parameters problem that can bias the remaining parameter estimates of interest. Koenker (2004) proposes a LASSO penalized estimation approach to deal with the incidental parameters problem associated with estimating many fixed effects in a nonlinear framework. 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)). In this framework, regression parameters are estimated as the

solution to the problem

$$\underset{\alpha, \beta, \gamma, \zeta \in \mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^n \sum_{t=1}^T w_{\tau} \rho_{\tau}(y_{it} - \alpha_i - \sum_{s=1}^k \zeta_s y_{i,t-s} - \mathbf{x}'_{i,t-1} \boldsymbol{\beta} - \mathbf{Z}'_t \boldsymbol{\gamma}) - \lambda \sum_i |\alpha_i| \quad (3.3)$$

where the weight w_{τ} controls the relative influence of the τ th 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 .

The standard approach to Bayesian quantile regression is based on the link between minimizing the errors of the frequentist quantile regression model and maximizing the likelihood of the asymmetric Laplace distribution (ALD) written as a scale mixture of normal distributions to leverage a simple and efficient Gibbs sampling algorithm (Yu and Moyeed (2001), Kozumi and Kobayashi (2011)). To augment the quantile regression model with adaptive LASSO shrinkage, a modification is made to the standard Bayesian implementation of LASSO shrinkage which places a Laplace prior on the parameter to be penalized (Park and Casella (2008)). The adaptive LASSO allows the LASSO penalty term λ_i shown in Equation (3.3) to vary across the individual effects. This allows larger estimated coefficients to be shrunk relatively less than smaller coefficients, so that the model can retain the most important unobserved individual heterogeneity while reducing the bias that may be imposed by selecting a single shrinkage parameter to be applied to each coefficient (Zou (2006)). Alhamzawi et al. (2012) and Aghamohammadi and Mohammadi (2017) implement the adaptive LASSO in a Bayesian quantile regression model via a Laplace prior placed on parameters to be penalized. In this case, the Laplace prior is placed on the individual effects $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$, written as

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

,

where the LASSO penalty term λ_i may vary across each individual effect α_i . Then, assuming $\nu_i = \lambda_i/\sqrt{\sigma}$, $i = 1, \dots, n$, the posterior distribution of $\boldsymbol{\alpha}$ can be written as

$$\begin{aligned} \pi(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\zeta}, \mathbf{y}, \mathbf{X}, \mathbf{Z}, \sigma, \lambda) &\propto \sigma^{-N} \nu^n \exp\left\{-\sum_{i=1}^n \sum_{t=1}^T \rho_{\tau} \frac{(y_{it} - \alpha_i - \sum_{s=1}^k \zeta_s y_{i \ t-s} - \mathbf{x}'_{i \ t-1} \boldsymbol{\beta} - \mathbf{Z}'_t \boldsymbol{\gamma})}{\sigma}\right\} \\ &\times \exp\left\{-\nu \sum_{i=1}^n |\alpha_i|\right\}. \end{aligned} \tag{3.4}$$

Considering σ and λ as nuisance parameters equates maximization of Equation (3.4) with respect to the model parameters to minimization of Equation (3.3). Also motivated by the ability to simplify estimation using a Gibbs sampling algorithm, Alhamzawi et al. (2012) implement the adaptive LASSO by writing the Laplace prior placed on the parameters to be penalized as a scale mixture of normal distributions. I follow Aghamohammadi and Mohammadi (2017), who propose a Bayesian hierarchical framework for panel data quantile regression with adaptive LASSO shrinkage that estimates Equation (3.3) via a tractable Gibbs sampling algorithm for posterior inference. Rewriting both the Laplace prior placed on the individual effects and the ALD placed on the likelihood function as scale mixtures of normal distributions suggests the following hierarchical model,

$$\begin{aligned}
y_{it}|e_{it}, \alpha_i, \beta, \zeta, \gamma, \sigma &\sim N(\alpha_i + \sum_{s=1}^k \zeta_s y_{i \ t-s} + \mathbf{x}'_{i \ t-1} \boldsymbol{\beta} + \mathbf{z}'_t \boldsymbol{\gamma} + (1 - 2\tau)e_{it}, 2\sigma e_{it}), \\
e_{it}|\sigma &\sim \text{Exp}(\frac{\tau(1 - \tau)}{\sigma}), \\
\boldsymbol{\beta}|\mathbf{b}_0, \mathbf{B}_0 &\sim N_j(\mathbf{b}_0, \mathbf{B}_0), \\
\boldsymbol{\gamma}|\mathbf{g}_0, \mathbf{G}_0 &\sim N_k(\mathbf{g}_0, \mathbf{G}_0), \\
\boldsymbol{\zeta}|\mathbf{q}_0, \mathbf{Q}_0 &\sim N_m(\mathbf{q}_0, \mathbf{Q}_0), \\
\alpha_i|s_i &\sim N(\alpha_i, s_i), \quad s_i|v_i^2 \sim \text{Exp}(\frac{v_i^2}{2}), \nu^2|\phi \sim \text{Exp}(\phi), \pi(\phi) \propto \frac{1}{\phi}, \\
\sigma|c_0, d_0 &\sim IG(c_0, d_0),
\end{aligned} \tag{3.5}$$

where $N(\cdot)$ represents a Normal distribution, $\text{Exp}(\cdot)$ represents an exponential distribution, and $IG(\cdot)$ represents the Inverse Gamma distribution.¹⁰ 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). 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. While the prior distributions of e , α , s , ν^2 , and ϕ are specifically motivated by implementing the Bayesian quantile regression with adaptive LASSO shrinkage in the hierarchical framework, I impose the following weak priors on the remaining parameters,

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

¹⁰See Aghamohammadi and Mohammadi (2017) for a full derivation of the model.

3.5 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.¹¹

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.¹² Second, this asset threshold partly deals with population changes due to bank asset growth and consolidation, as well as secular trends associated with inflation.

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)). Similar to Covas et al. (2014), I estimate 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

¹¹The 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 Banks in 1997,” Federal Reserve Bulletin, vol. 84 (June), p. 408.

¹²The 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>

the effect of the bias is insignificant, since the initial conditions will have little impact on parameter estimates.¹³

The remaining bank 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. Table 3.1 presents summary statistics for all of the variables used in the empirical analysis.

3.6 Estimation Results

3.6.1 Full Sample Results

As a first point of comparison, I estimate fixed-effects OLS (FE-OLS) parameters of model 3.1 in addition to the BPQR-AL parameter estimates. Table 3.2 shows the FE-OLS and BPQR-AL results, estimated across all banks in the sample over the period from 1998 Q1-

¹³In fact, imposing this filter on my sample results in the average bank being observed for roughly 20 years.

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 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.¹⁴ 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

¹⁴Trace and ACF plots were used to monitor the convergence of the Gibbs sampling algorithm, and suggest satisfactory convergence of the model.

tails of the distribution, while the conditional median of the distribution is more strongly influenced by past realizations.

Figure 3.5 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.

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 realized 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.

3.6.2 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 et al. (2023), Laliberte

and Sengupta (2024), Sengupta and Xue (2022)). 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 3.6 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. This result suggests that bank NIMs in the conditional tails of the distribution became relative more exposed to interest rates in the post-2020 period.

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.

3.6.3 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.¹⁵ 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. However, a reasonable approach is 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 3.7 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

¹⁵In the FR Y-9c form, expenses on deposits are broken into (i) time deposits, and (ii) other deposits. This captures interest expenses on all non-time interest-bearing deposits, according to the form instructions. I assume these are primarily expenses on interest-bearing demand deposits. I do not include expenses paid on foreign deposits in any categories.

quantile-specific estimates for the component of interest income in the two subsample time periods.¹⁶ The figure shows that interest income on loans as a share of interest-earning assets explains the largest increase in sensitivity to the 3-month Treasury and slope of the yield curve compared 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 3.8 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 line with the general view in the academic literature that demand deposit expenses are priced with consideration to short-term interest rates. 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.

The level of demand deposits' sensitivities to the 3-month Treasury rate is notably smaller than those observed for loan income. While the smaller level may be driven in part by banks' ability to reprice deposits less than changes in market interest rates, the time variation in deposit sensitivities may also be constrained when interest expenses approach the zero lower bound. Figure 3.9 shows the model results for loans and demand deposits, with the inclusion of an additional time subsample spanning 1998-2021, prior to the Fed Funds rate increase beginning in 2022. This figure shows that the increased lower tail interest rate sensitivities associated with loans had already occurred by 2021 while the increased upper tails interest rate sensitivities for demand deposits did not occur until 2022, after short-term interest rates increased. This finding suggests that, during the low interest rate period between 2020-2021,

¹⁶Given the focus on time variation in interest rate parameter estimates, I do not show the credit risk or bank balance sheet parameter estimates, although these covariates were included in the estimations.

increased loan sensitivities to interest rates drove lower conditional tails of NIM while banks deposit expenses may have already been constrained at near minimum levels.

3.6.4 Results by Bank Size

Studies have documented NIM differentials in terms of bank size (Abdymomunov et al. (2023), Sengupta and Xue (2022)). 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 3.10 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)).

Given the time variation found in interest rate sensitivities and the heterogeneity in NIM across the bank size groups, it is useful to examine how the time variation changes across the size distribution. Figure 3.11 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.

3.7 Density Forecasting

When examining bank NIM exposure to interest rate risk, it is desirable for regulators to select a model that most accurately describes conditional outcomes. Given the relative complexity of the BPRQ-AL model to a simpler alternative such as a linear model, it is worth understanding if there are benefits given the costs of model implementation. 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 simpler 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, quantile-varying individual effects and adaptive LASSO shrinkage may be better suited to capturing the tails of the conditional distribution relative to the conditional mean estimation, or variants of the BPQR-AL model that do not deal with the incidental parameters issue or impose more strict assumptions on the LASSO shrinkage.

I evaluate the density forecast performance of the BPQR-AL model relative to several benchmark models that capture various aspects of the BPQR-AL model. To capture individual effects and heterogeneous sensitivities, I include fixed-effect OLS (FE-OLS) and pooled quantile regression (QR) benchmark specifications of equation (3.1). Similar to Covas et al. (2014), I also estimate a frequentist fixed-effects quantile autoregressive model (FE-QAR) specification based on equation (3.3), which imposes a single LASSO shrinkage parameter across quantile-invariant individual effects.¹⁷ In each case, I use the same set of model covariates reported in the previous estimation results. If heterogeneous sensitivities and individual effects in the tails of the conditional distribution matter, and if adaptive LASSO shrinkage

¹⁷In the FE-QAR model estimation, I follow Covas et al. (2014) in setting the single shrinkage parameter λ equal to 1.

on the individual effects improves issues of biased parameter estimates and the incidental parameter problem relative to the regular LASSO, then forecasts of the BPQR-AL model should outperform FE-OLS, QR, and FE-QAR methods.

3.7.1 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. 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.¹⁸ 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 et al. (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,

$$f_y(\mu, \sigma, \alpha, \nu) = \frac{2}{\sigma} t\left(\frac{y - \mu}{\sigma}; \nu\right) T\left(\alpha \frac{y - \mu}{\sigma} \sqrt{\frac{\nu + 1}{\nu + \left(\frac{y - \mu}{\sigma}\right)^2}}\right), \quad (3.6)$$

where $t(\cdot)$ and $T(\cdot)$ represent the probability density and cumulative distribution func-

¹⁸While 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.

tions 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 et al. (2019), this is equivalent to an exactly identified 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.

3.7.2 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

$\{1, 2, \dots, P\}$, given a quantile forecast f at time t for bank i , the QS is defined as

$$QS_{it}(\tau_p) = \begin{cases} (1 - \tau_p)(f_{it}(\tau_p) - y_{it}) & \text{if } y_{it} < f_{it}(\tau_p) \\ \tau(y_{it} - f_{it}(\tau_p)) & \text{if } y_{it} \geq f_{it}(\tau_p) \end{cases} \quad (3.7)$$

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

$$CRPS_{it} = \frac{2}{P} \sum_{p=1}^P QS_{it}(\tau_p) \quad (3.8)$$

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

$$qwCRPS_{it} = \frac{2}{P} \sum_{p=1}^P \gamma(\tau_p) QS_{it}(\tau_p), \quad (3.9)$$

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 $\{0.005, 0.01, 0.015, \dots, 0.995\}$ 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 produces a more (less) accurate density forecast 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.¹⁹

Table 3.3 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 have an improved CRPS point estimate 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.

¹⁹Despite the relatively short number of periods in the out-of-sample period, the panel data allow for greater statistical power in conducting this test.

3.8 Measures of 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 more adverse tail risks. Figure 3.12 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.

Aggregating across the individual bank density forecasts may also be useful in a macro prudential setting where regulatory concerns may lie around aggregate risks rather than bank-specific risks. 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 across the quantiles of each individual distribution (Busetti (2017)). 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 , $i = 1, \dots, n$. 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. For simplicity, I set the ω_i such that the conditional quantiles are equally weighted.

Figure 3.13 shows the aggregated density forecasts along with the average realization

of NIM across all banks in the sample at each time period across the FE-OLS, QR, and FE-QAR models for comparison. 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 but lower probabilities 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. 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.

Finally, one benefit of a model framework that recovers the conditional distribution is the ability to construct measures of tail risk, also useful to banking regulators. One way to measure tail risks associated with the conditional density forecasts produced by the BPQR-AL model is the expected shortfall approach. For a chosen probability π , the expected NIM shortfall SF for bank i at forecast horizon $t + 1$ is defined as

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

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

I follow Giglio et al. (2021) in computing the cross-sectional dispersion of expected shortfall across banks and over time at the 5 percent level, setting $\pi = 0.05$. To place economic context on the shortfall measure, I compute the shortfall in terms of bank bottom-line profitability, return on average assets (ROA). ROA is defined as the total net income of a bank divided by its average total assets, and bank NIM is typically an important component of

bank net income.²⁰ Net income is important for bank regulatory capital because retained earnings is a primary way that banks can build loss-absorbing regulatory capital, suggesting that it is important for financial stability. I estimate ROA conditional on the expected NIM shortfall measure by multiplying the expected shortfall by interest-earning assets to estimate the shortfall in dollars, and then subtract the difference between the realized dollar value and the shortfall dollar value from bank total net income.

The left panel of Figure 3.14 shows the distribution of ROA for bank holding companies in the sample from 1998-2022 for the 5th, 25th, 50th, 75th, and 95th percentiles of the distribution. In benign periods the average ROA is roughly 1.5 percent of average total assets, however the extreme tails extended to less than negative 8 percent of total assets during the GFC period. 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 decreased bank NIM and increased provision expenses.²¹ The right panel shows the same percentiles for the cross-sectional distribution of ROA that now including the 5 percent expected shortfall change in NIM for all banks in the sample. The magnitude of ROA downside risks suggested by the expected shortfall measure peaks in the GFC period and declines afterward until a sharp increase in downside risk beginning in 2020. The shortfall is sizable enough during the GFC and post-2020 periods such that banks within the interquartile range extend near or into negative ROA levels, suggesting that tail risks to bank NIM are material to overall bank profitability.

3.9 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

²⁰Other components of ROA include noninterest income and expenses, provisions, realized gains and losses, taxes, and other adjustments.

²¹See <https://www.federalreserve.gov/publications/2020-november-supervision-and-regulation-report-bank.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. 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. The analysis examines a relatively large sample of US bank holding companies with at least \$500 million in total assets from 1998-2022, and utilizes a novel Bayesian dynamic panel quantile regression that deals with the problem of incidental parameter bias present in nonlinear estimation of panel data. I 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, with zero lower bound constraints impacting tail outcomes for bank NIMs by constraining pass-through of short term interest rate changes to deposits. 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 et al. (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 observed and unobserved heterogeneity at the bank level estimated by quantile regressions with individual effects offer a more accurate picture of

potential outcomes relative to conditional mean specifications. Density forecasts also allow for computation of expected shortfall 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. 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.

3.10 Tables

Table 3.1: Estimation Sample Summary Statistics, 1998-2022.

This table shows summary statistics for data used in the model estimation. The 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 et al. (2018). Source: US FRY9-c and Call Report Filings, H.15 interest rate data, ICE Data Indices, LLC

Dependent Variable								
Variable	n	Mean	SD	Min	P25	P50	P75	Max
Financial Variables								
NIM	24326	3.60	0.93	-12.85	3.13	3.60	4.09	10.35
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
Bank Balance Sheet Variables								
BBB-10y Spread	24326	1.59	0.83	0.72	1.10	1.42	1.81	5.80
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 3.2: FE-OLS and BPQR-AL Model Parameter Estimates, 1998-2022.

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 et al. (2018).

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

Table 3.3: CRPS Ratio of BPQR-AL Density Forecasts to Alternative Models.

This table shows ratios of average continuous ranked probability score (CRPS) and quantile-weighted CRPS of the BPQR-AL model relative to alternative model specifications. Values less (greater) than 1 indicate the BPQR-AL model produces a smaller (larger) CRPS relative to the comparison model. 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). Left tail, center, and right tail-weighted CRPS are denoted as CRPS-L, CRPS-C, and CRPS-R, respectively. 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.

3.11 Figures



Figure 3.1: Average Annualized Bank Net Interest Margin, Selected Interest Rates 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.

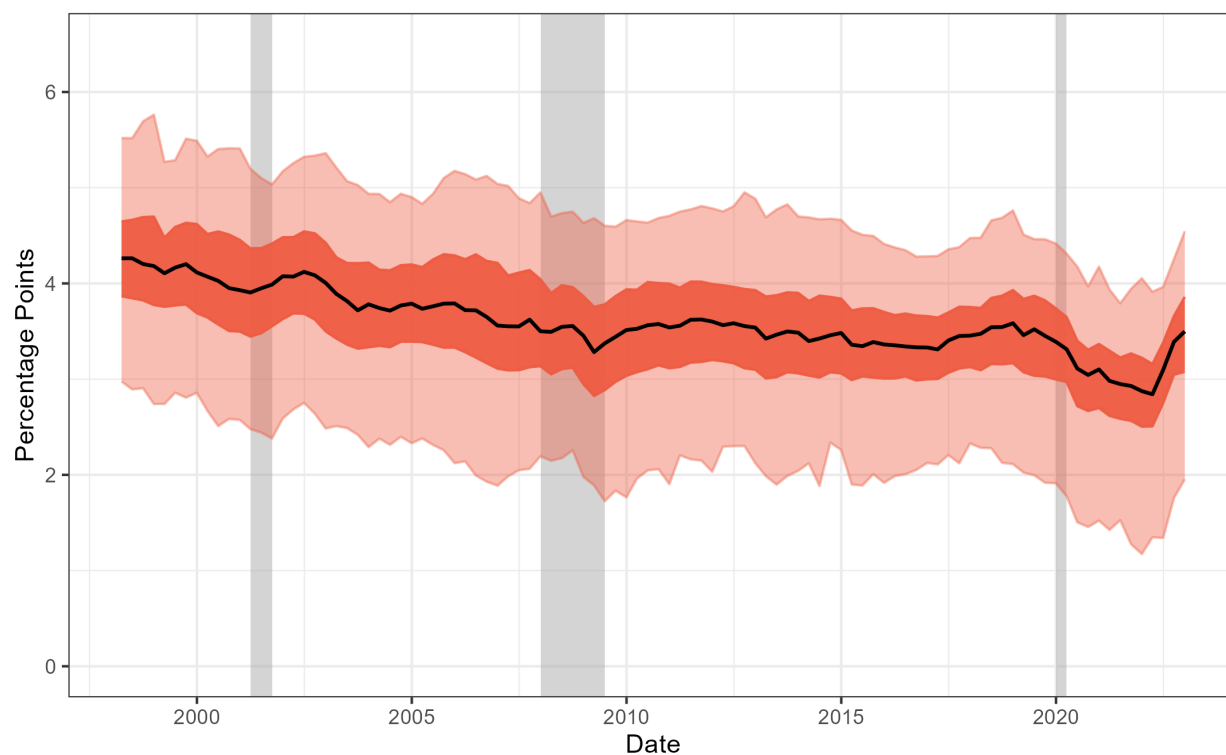


Figure 3.2: Distribution of Annualized NIM, 1998-2022.

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.

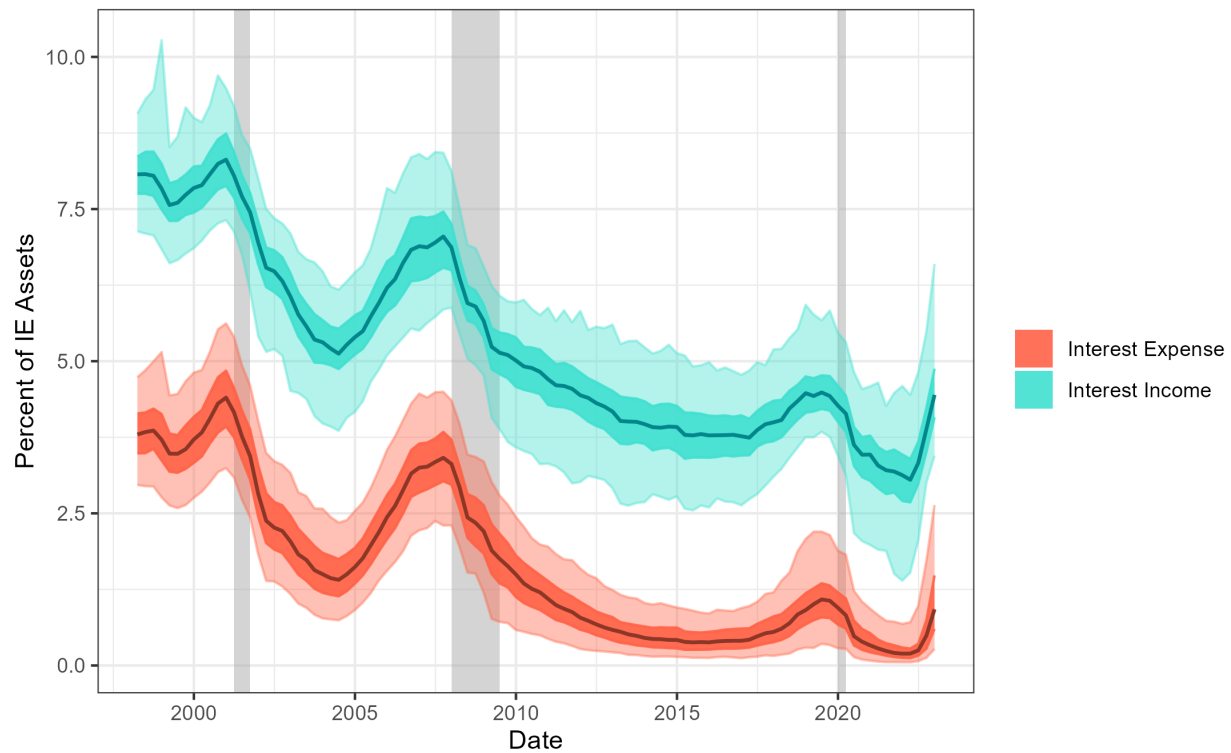


Figure 3.3: Distribution of Interest Income and Interest Expense as a Percent of Interest-Earning Assets, 1998-2022.

Shaded areas represent 5th, 25th, 75th, and 95th percentile cross-sectional outcomes, and solid lines indicate cross-sectional medians. 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).

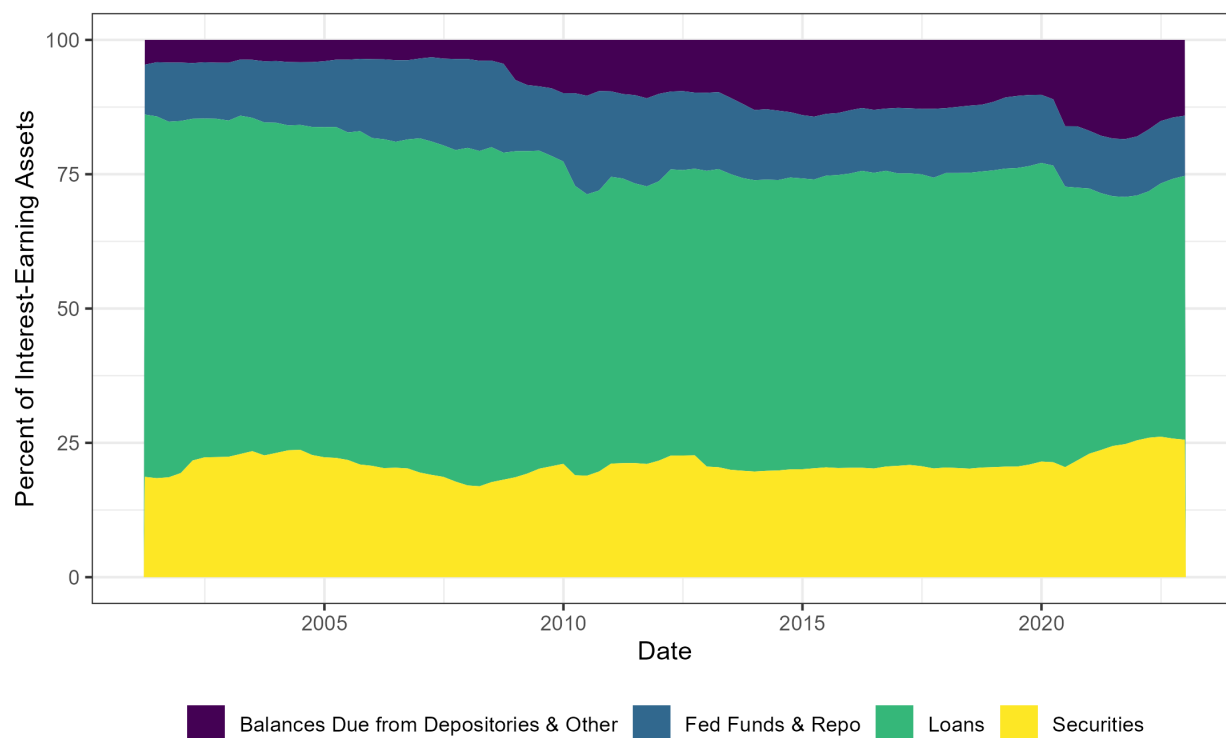


Figure 3.4: Aggregate Components of Interest-Earning Assets as a Percent of Total Interest-Earning Assets, 2001-2022.

Data from the period 1998-2000 are excluded from the 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)

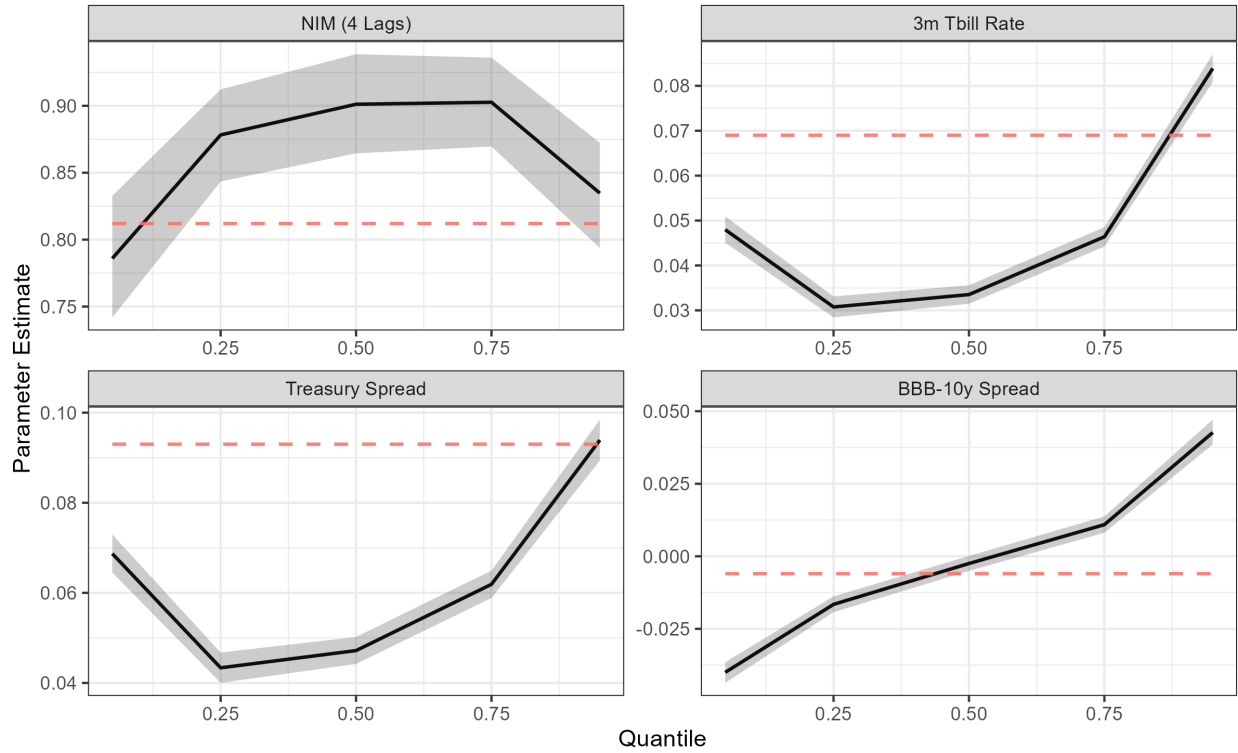


Figure 3.5: BPQR-AL Quantile Parameter Estimates, 1998-2022. BPQR-AL model estimated at 5th, 25th, 50th, 75th and 95th conditional quantiles.

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. 3m Tbill Rate refers to the 3-month Treasury bill rate. BBB-10y Spread refers to the BBB spread relative to the 10-year Treasury rate. 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 et al. (2018).

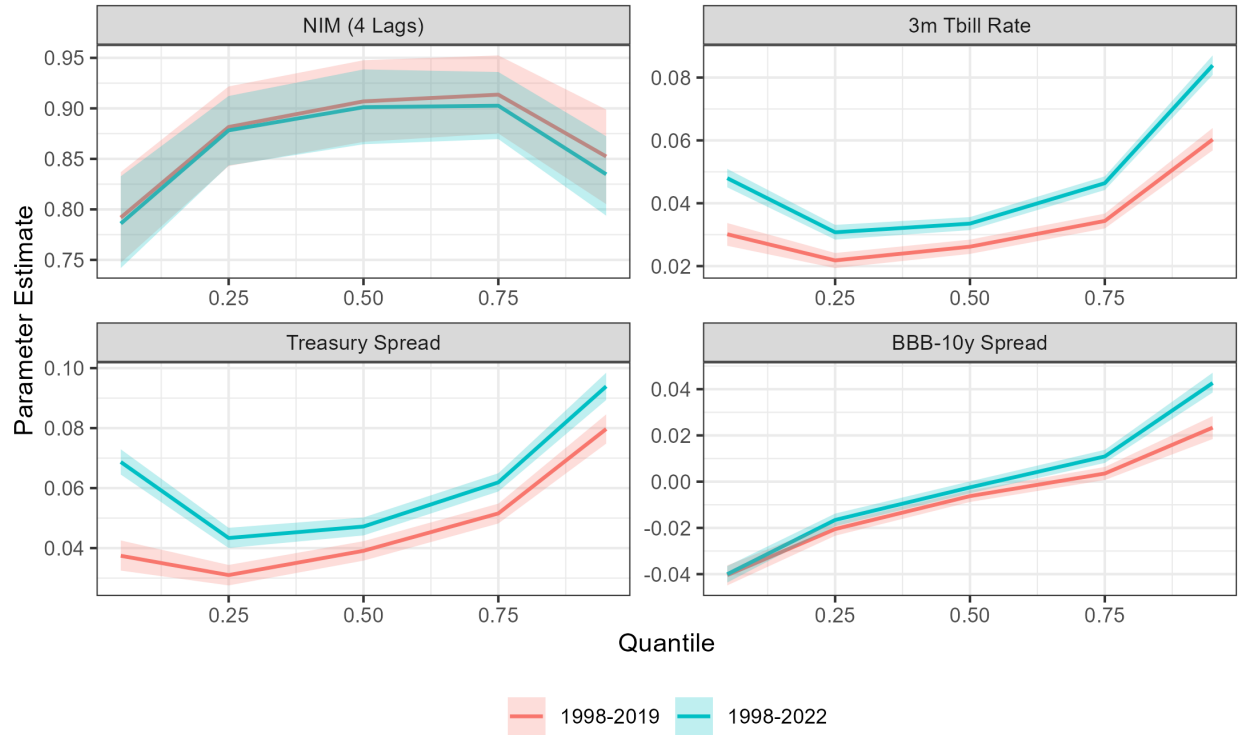


Figure 3.6: BPQR-AL Quantile Parameter Estimates by Time Period Subsamples Spanning 1998-2019 and 1998-2022.

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. 3m Tbill Rate refers to the 3-month Treasury bill rate. BBB-10y Spread refers to the BBB spread relative to the 10-year Treasury rate.

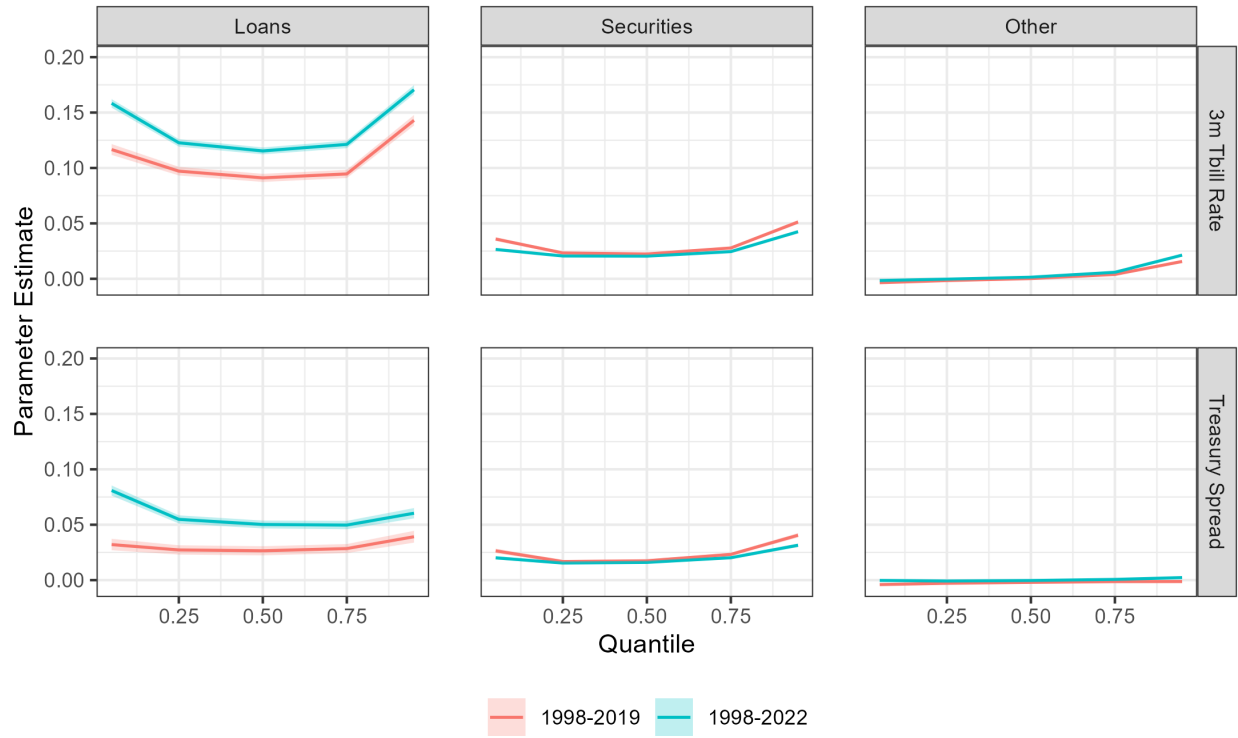


Figure 3.7: BPQR-AL Quantile Parameter Estimates by Time Period subsamples Spanning 1998-2019 and 1998-2022.

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. 3m Tbill Rate refers to the 3-month Treasury bill rate. BBB-10y Spread refers to the BBB spread relative to the 10-year Treasury rate.

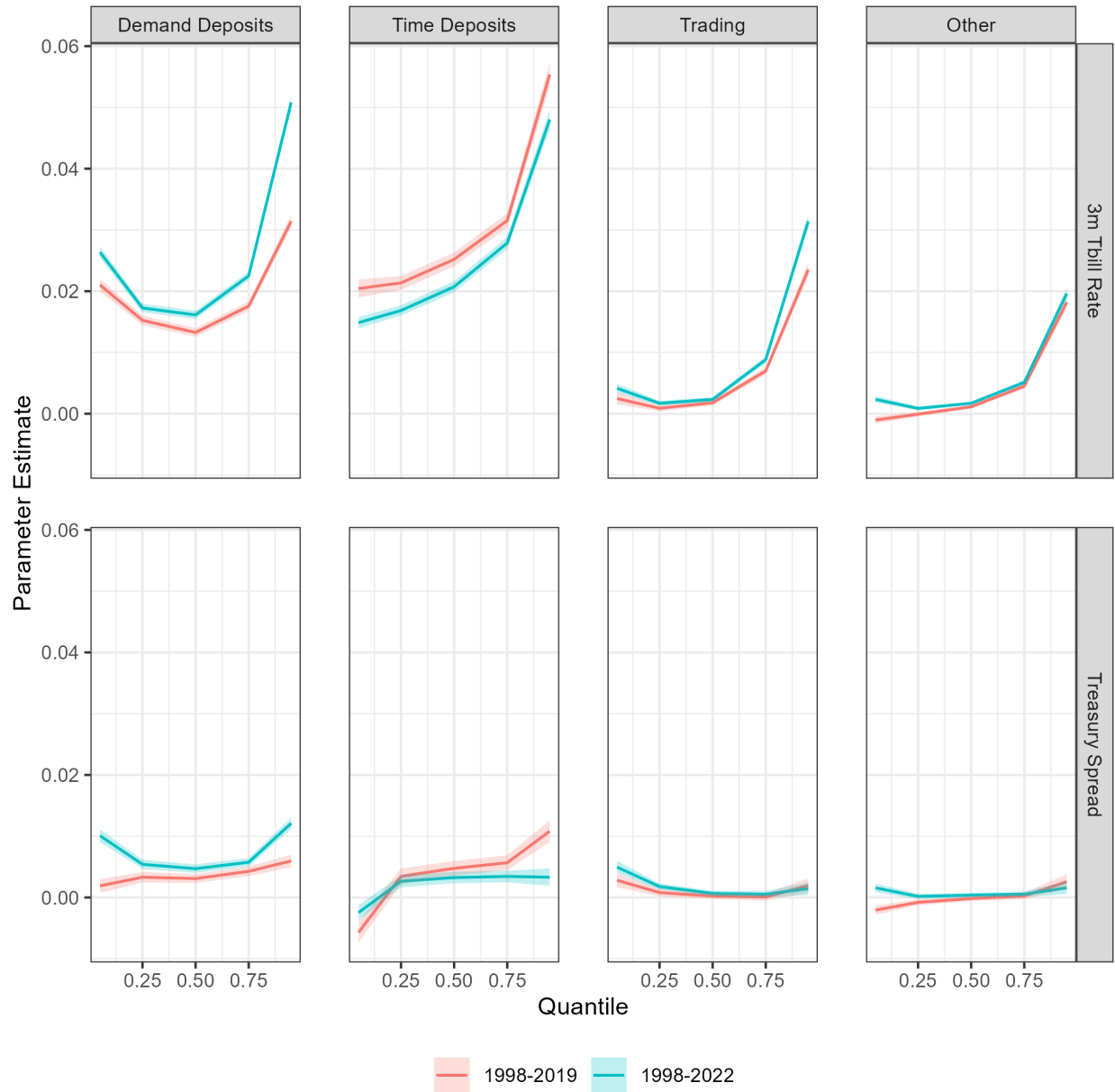


Figure 3.8: BPQR-AL Quantile Parameter Estimates by Time Period Subsamples Spanning 1998-2019, and 1998-2022.

The 'other' category of interest expenses includes subordinated notes and debentures, fed funds and repo, and other. 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. 3m Tbill Rate refers to the 3-month Treasury bill rate.

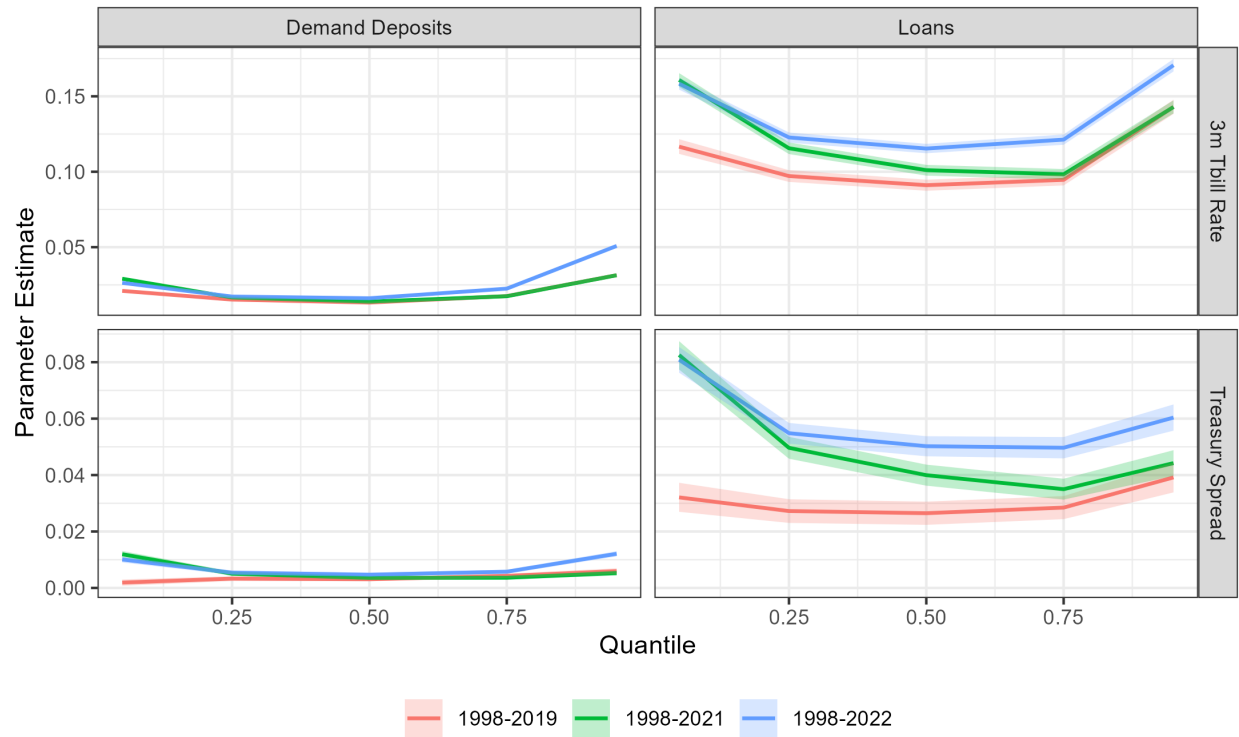


Figure 3.9: BPQR-AL Quantile Parameter Estimates for Demand Deposits and Loans by Time Period Subsamples Spanning 1998-2019, 1998-2021, and 1998-2022.

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. Solid lines represent posterior mean estimates, and shaded areas represent 95% credible intervals. 3m Tbill Rate refers to the 3-month Treasury bill rate.

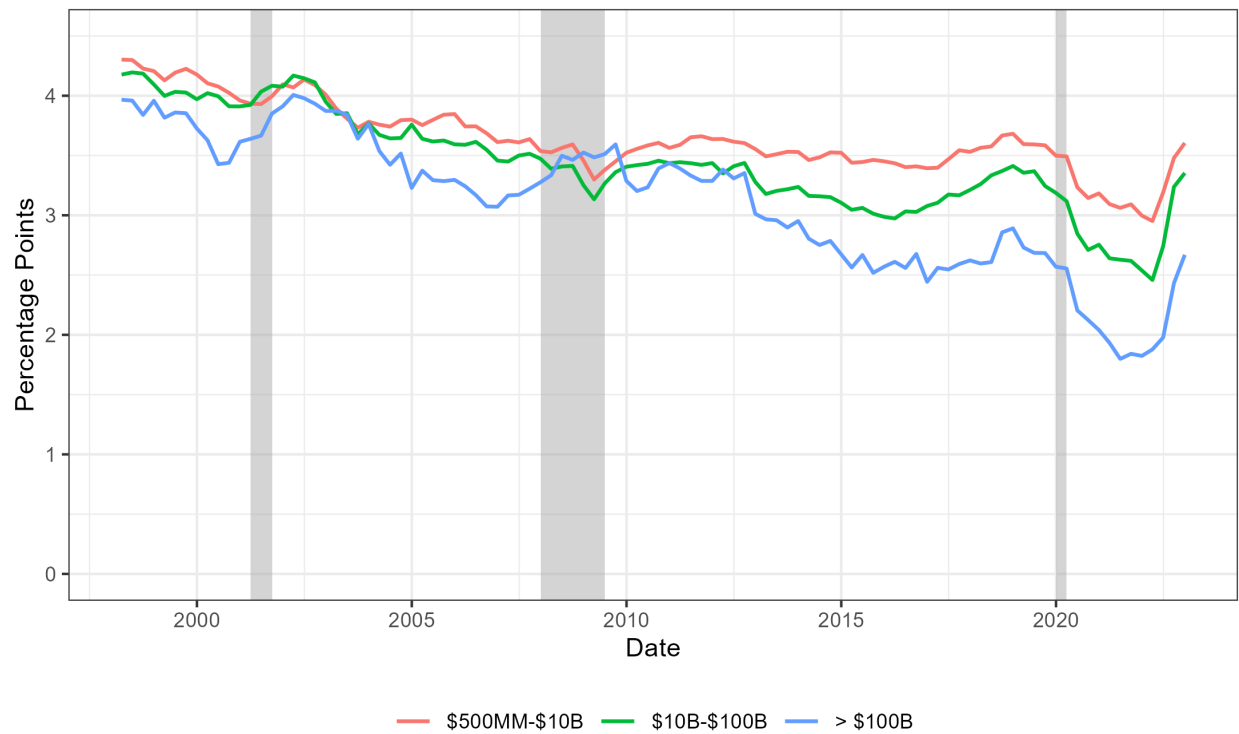


Figure 3.10: Average Annualized Net Interest Margin, by Size, 1998-2022.
 Vertical grey areas represent NBER-defined US recessions. Source: Merger-Adjusted FRY9-c Data (MAY9c)

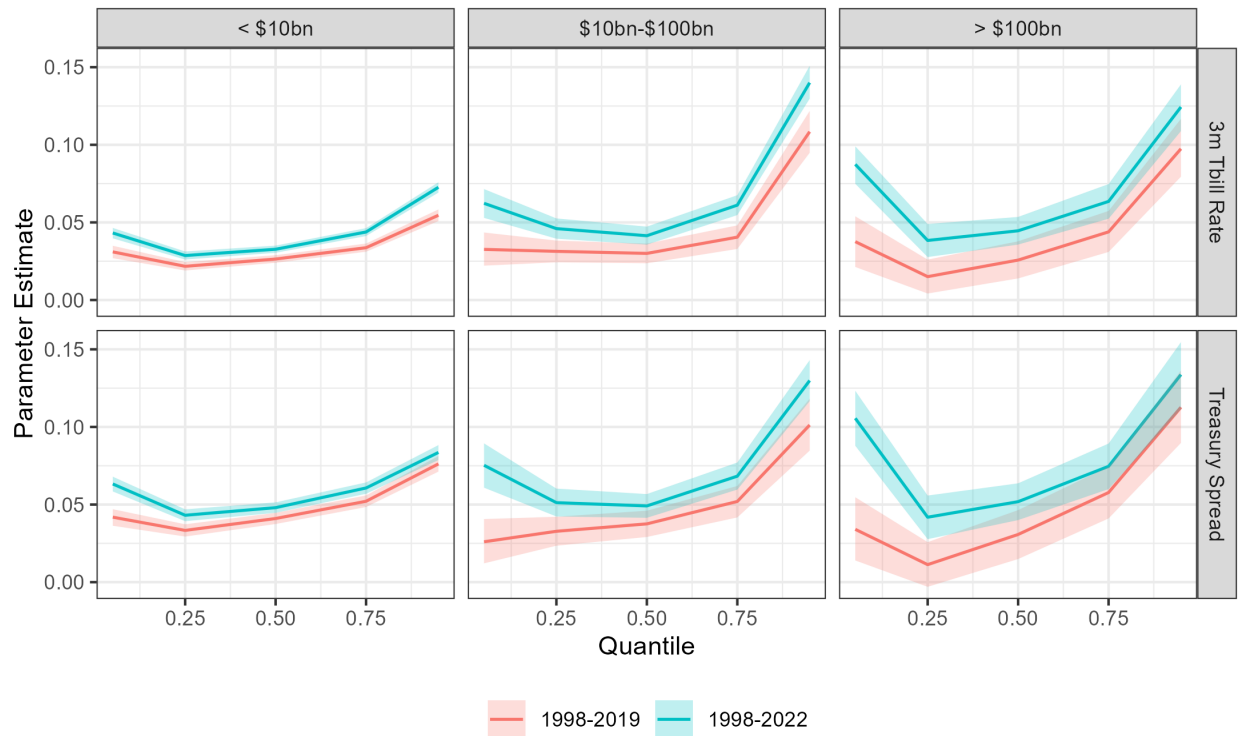


Figure 3.11: BPQR-AL Interest rate factor quantile parameter estimates by bank size group and time period.

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. 3m Tbill Rate refers to the 3-month Treasury bill rate. BBB-10y Spread refers to the BBB spread relative to the 10-year Treasury rate.

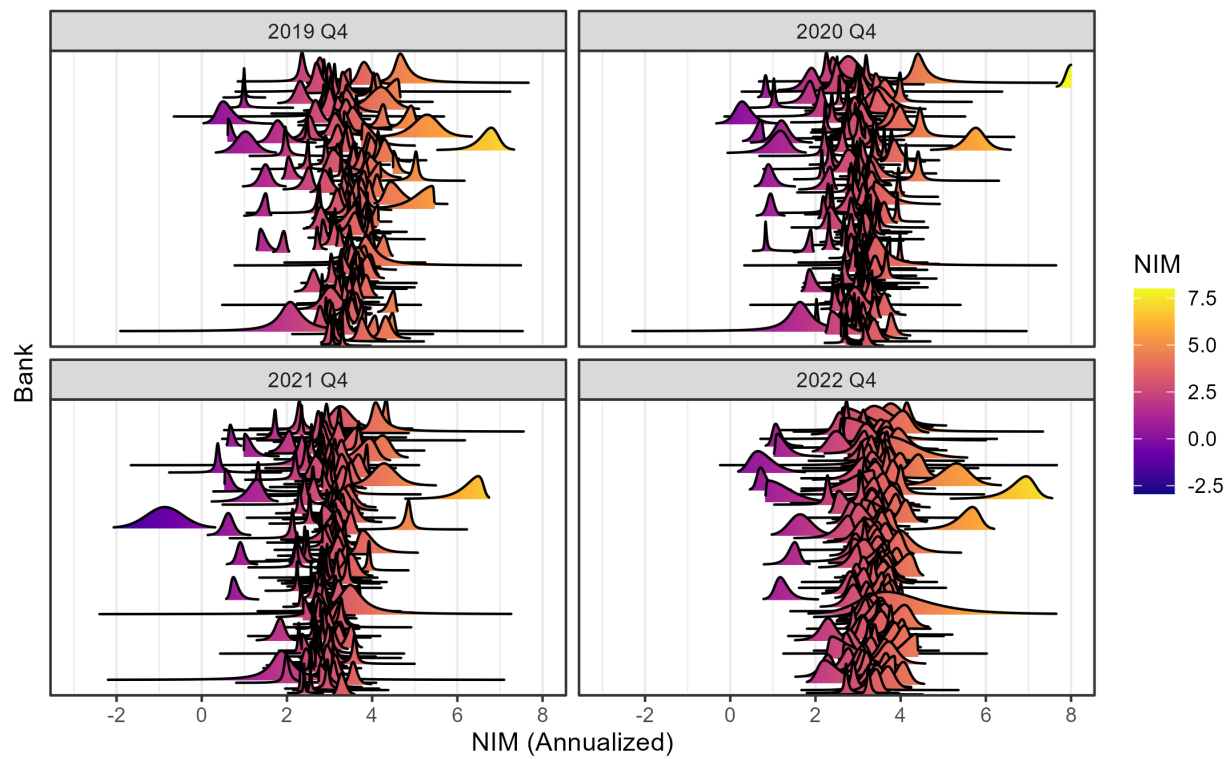


Figure 3.12: Individual bank density forecasts based on BPQR-AL pseudo out-of-sample estimations

Density forecasts are estimated from the BPQR-AL quantile forecasts fitted to the skewed-t distribution proposed by Azzalini and Capitanio (2003).

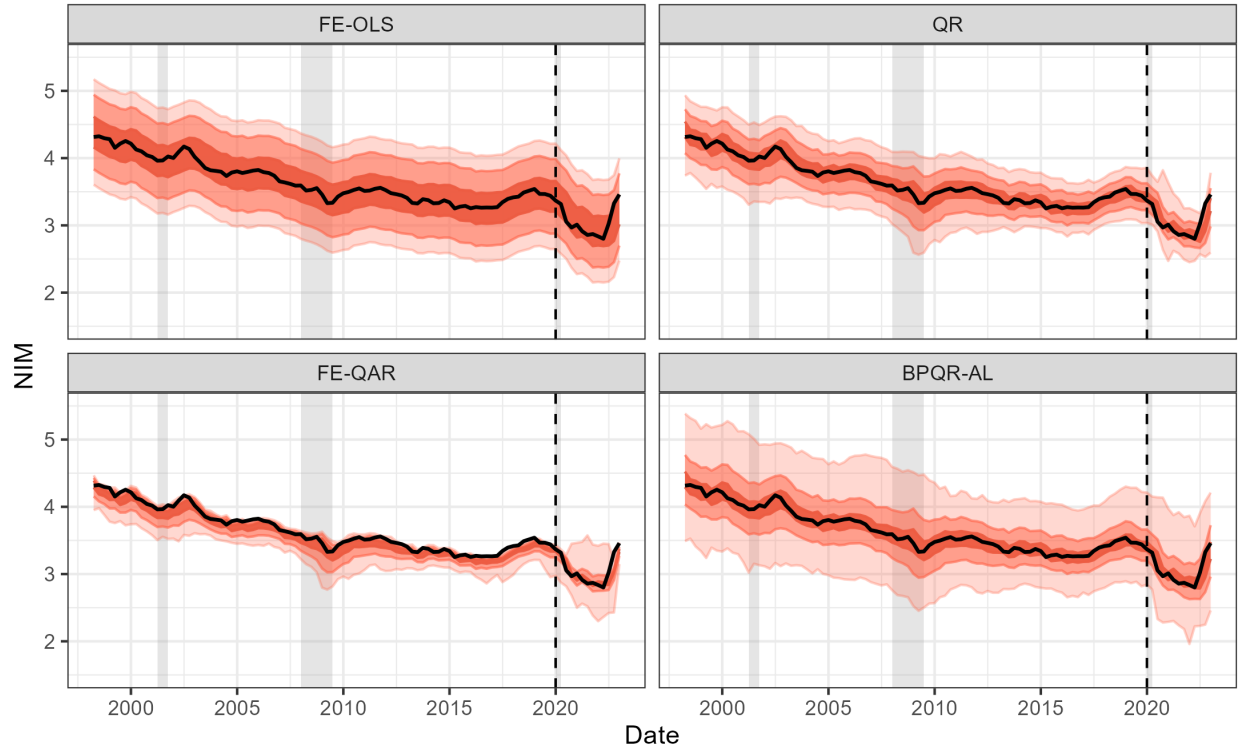


Figure 3.13: Aggregate NIM density forecasts by model.

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 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.

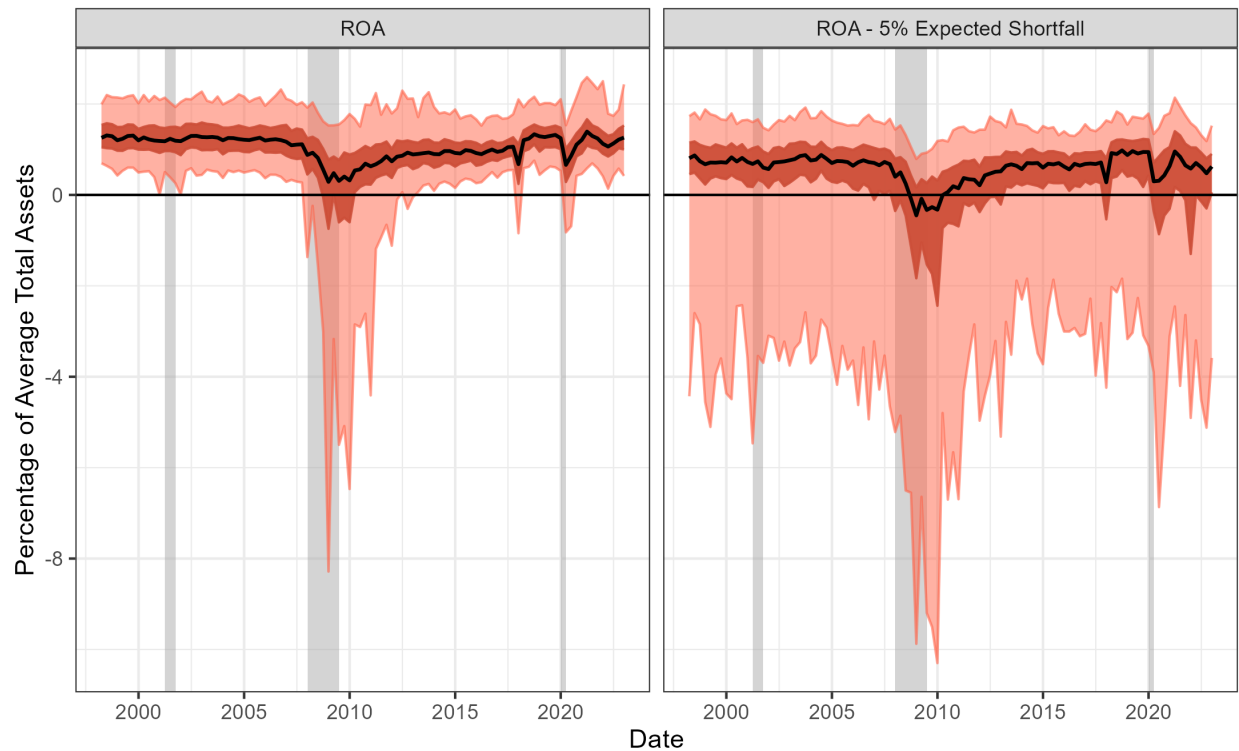


Figure 3.14: Actual Return on Assets (left panel) and Return on Assets Under 5 Percent Expected NIM Shortfall (right panel).

The expected shortfall is computed from BPQR-AL forecasts (right panel) across bank holding companies, 1998-2022. 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 5th, 25th, 75th, 95th percentiles, solid black line represents the median. Density forecasts are estimated from the BPQR-AL quantile forecasts fitted to the skewed-t distribution proposed by Azzalini and Capitanio (2003)

3.12 Internet Appendix

3.12.1 Sample Attrition

The data used in this paper rely on US bank holding company filings of form FR Y-9c. Regulatory filing requirement changes that occurred during the sample period result in sample attrition, causing some banks to exit the sample as a result. 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 driven by inflation, consolidation, growth of asset balances, and reduction of bank burden.²² 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.

3.12.2 Adaptive LASSO

To demonstrate the impact of the adaptive LASSO in the BPQR-AL, Figure 3.15 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

²²See https://www.federalreserve.gov/apps/reportingforms/Report/Index/FR_Y-9C.

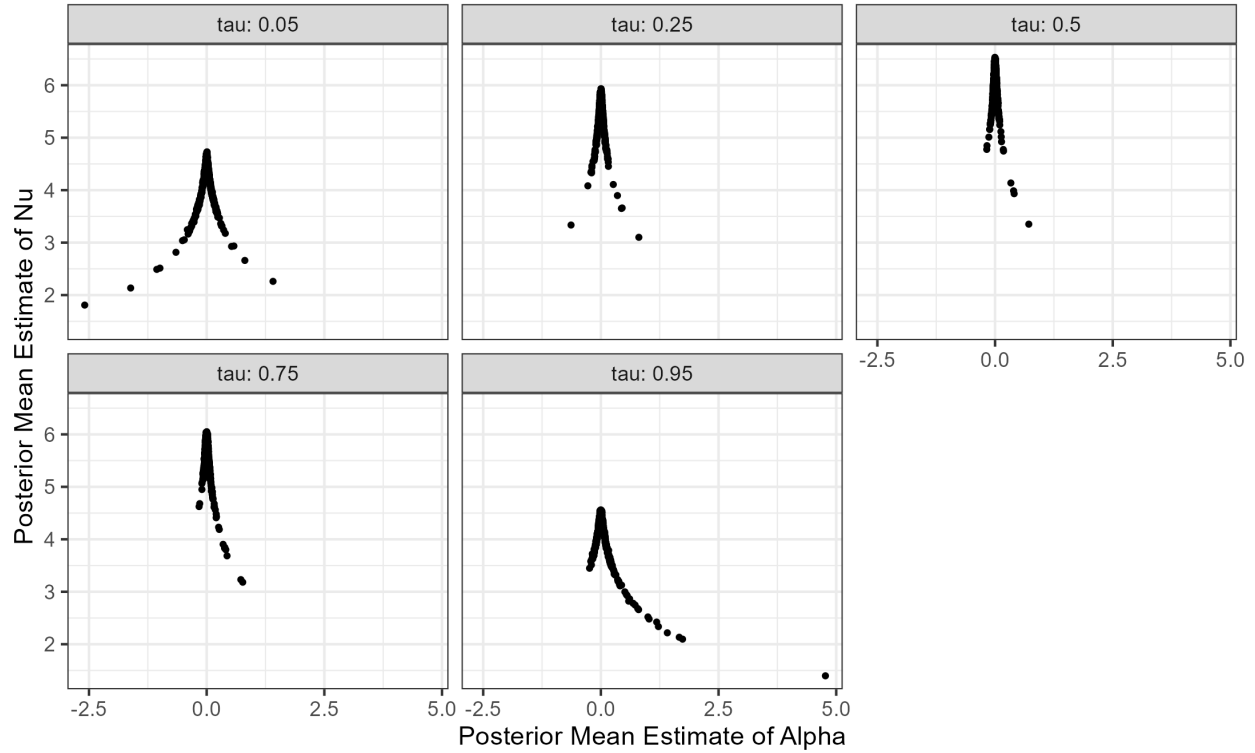


Figure 3.15: BPQR-AL Individual Effects Relative to Shrinkage Parameter (ν), 1998-2022.

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.

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.²³ 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.

²³There 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).

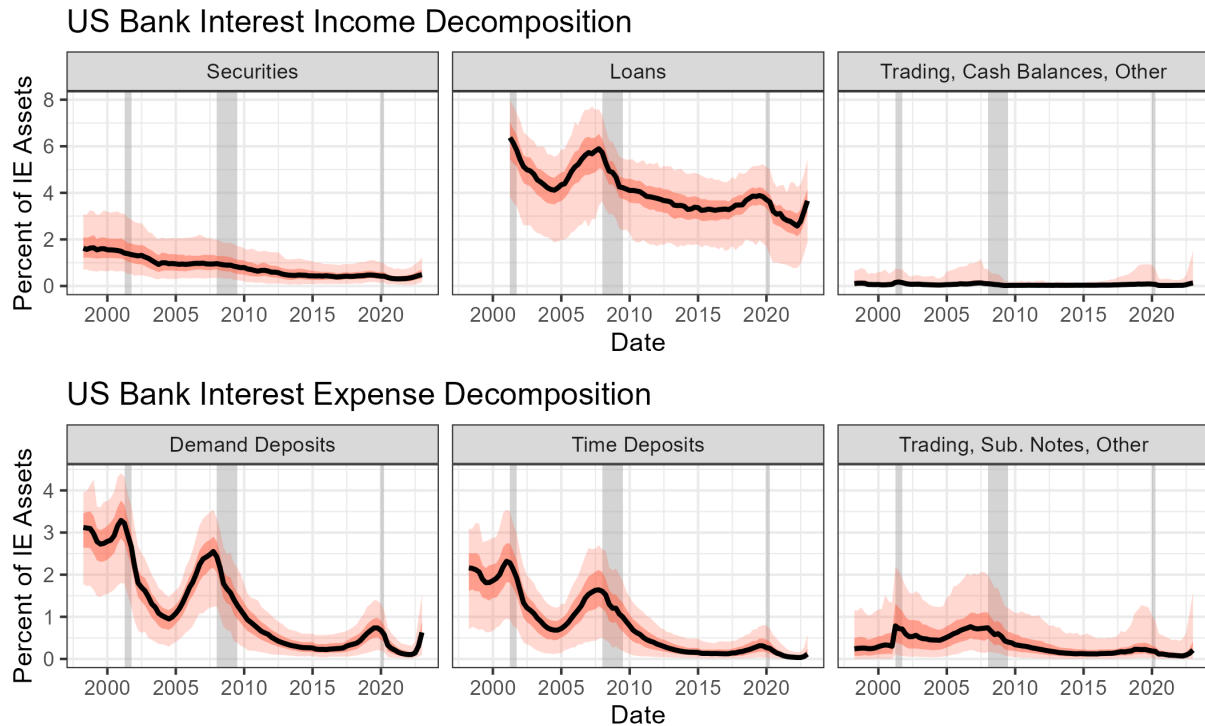


Figure 3.16: Income and Expense decomposition of Cross-Sectional Net Interest Margins, 1998-2022.

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).

3.12.3 NIM Numerator Decomposition

Figure 3.16 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.



Figure 3.17: 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).

3.12.4 Loan and Security Concentrations in the Cross-Section

Figure 3.17 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, 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.

3.12.5 Bank Size Results

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. Figure 3.18 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 lower at the higher tail of the distribution. This being said, the 95 percent credible interval around estimates for the largest banks are relatively wide for the sum of autoregressive terms, likely a result of the small number of banks in the largest size group. 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. A 100 basis point increase in both the 3-month Treasury bill rate and the Treasury spread is associated with an additional 4 basis 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 be relatively more willing and able to fund riskier projects. 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.

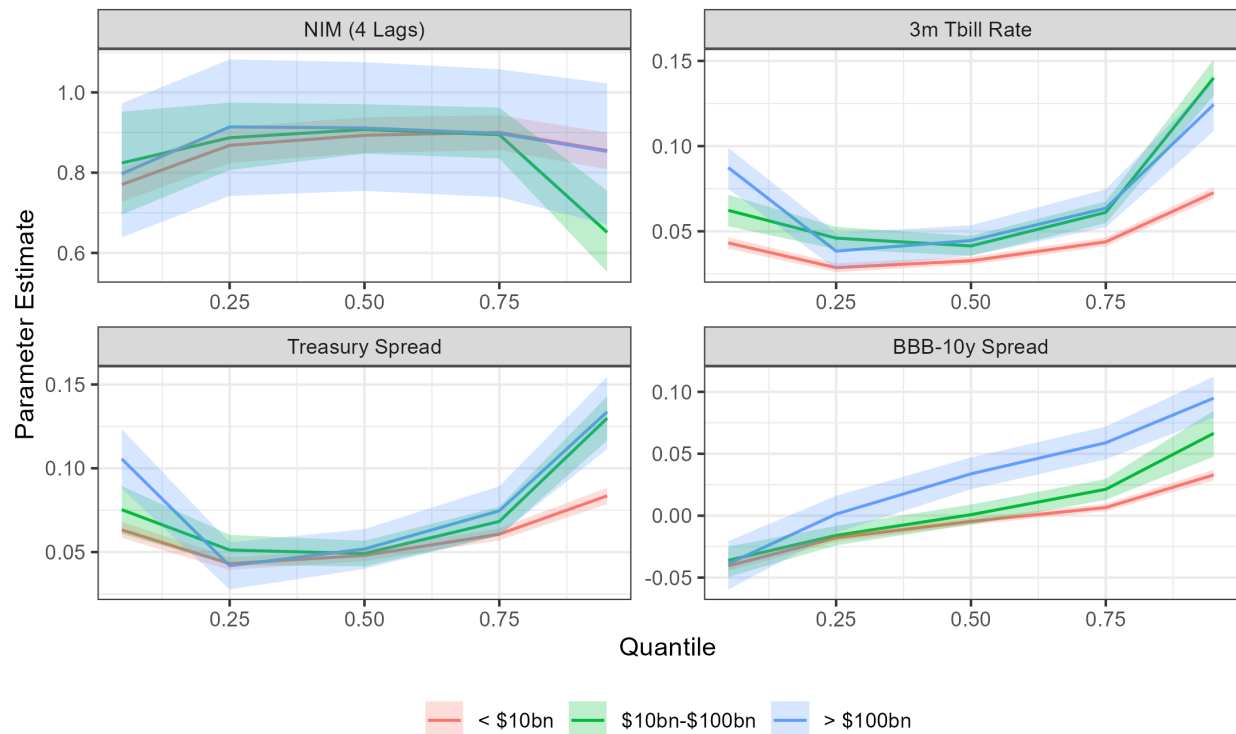


Figure 3.18: BPQR-AL Parameter Estimates by Bank Size Group, 1998-2022.

Bank size groupings are 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.

3.12.6 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 3.12.5. 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.

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 (3.1) 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 3.4 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.

Table 3.4: CRPS Tests of Macroeconomic Factor Importance.

This table shows 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.

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 β and γ , and scale parameter σ given 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 β and γ in the hierarchical specification of (3.5). 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 et al. (2024) and Carriero et al. (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.²⁴ The hyperparameters λ_1 and λ_2 are set equal to 0.2, following Carriero et al. (2020).

Table 3.5 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

²⁴For 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.

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.

Table 3.5: CRPS Tests of Prior Sensitivity.

This table shows 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.

Bank Size Subsample Models

Parameter results shown in Section 3.12.5 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 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 3.6 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.

Table 3.6: CRPS Tests of Bank Size-Specific Models.

This table shows 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.

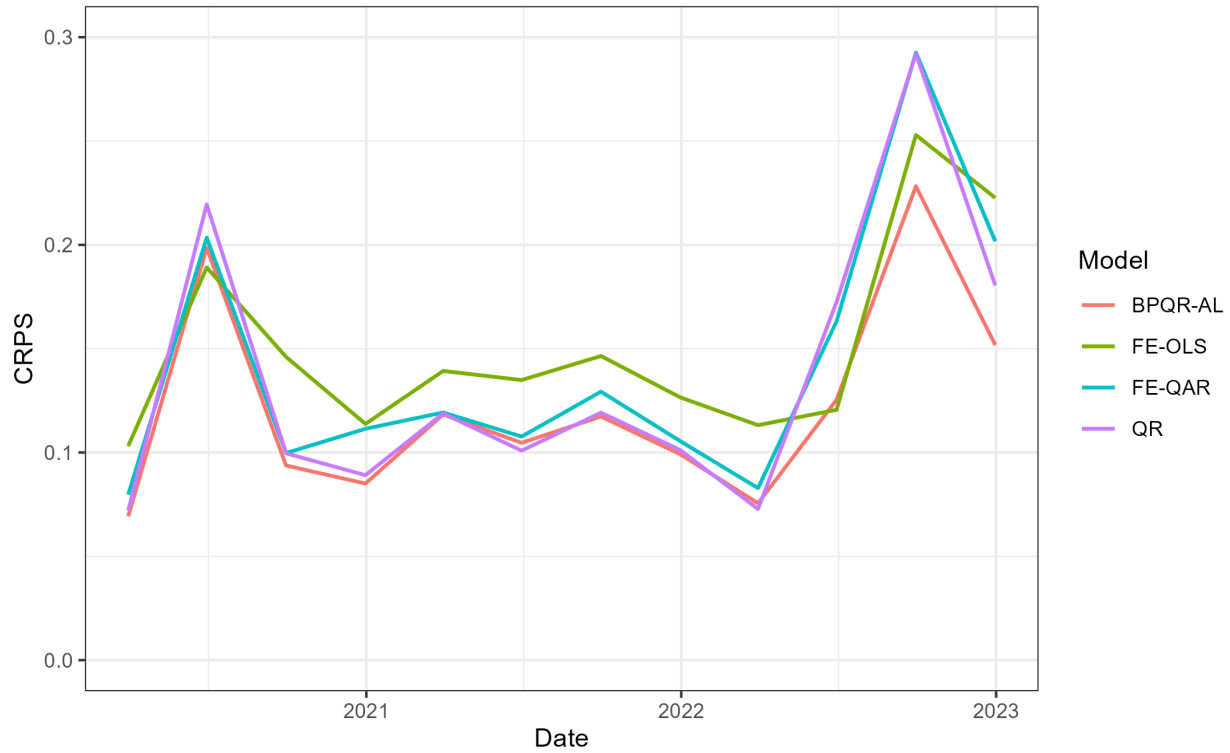


Figure 3.19: Pseudo Out-of-Sample CRPS by Time Period Across Alternative Models.

The set of models includes fixed-effects OLS (FE-OLS), pooled quantile regression (QR), Fixed-Effects Quantile Autoregressive model (FE-QAR) and Bayesian panel quantile regression with adaptive LASSO (BPQR-AL). Pseudo out-of-sample period spans 2020-2022.

3.12.7 Alternative Model Forecast Performance Over Time

Figure 3.19 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.

Chapter 4

Macroeconomic Parameter Instability in Auto Loan Loss Models

4.1 Introduction

This paper estimates a discrete-time Markov transition model of auto loan performance using multinomial logistic regressions. The model is estimated over the 2000-2025 period, which includes the global financial crisis (GFC), and the Covid shock. The first purpose of this paper is to evaluate the stability of estimated coefficients on macroeconomic shocks like unemployment. These coefficients matter for the stress testing models that banking regulatory agencies run to evaluate bank capital and they matter for loss forecasting by auto lenders. Using rolling regressions we demonstrate that these coefficients change significantly in response to the large aggregate shocks of the GFC and the Covid periods and the effect on these estimates is persistent.

Figure [4.1](#) shows the instability of the coefficient on our unemployment variable for rolling regression estimates of the relative risks of a current loans going into default. The effects are particularly evident for the Covid period, when despite large spikes in unemployment, auto lending default rates declined. While not necessarily surprising due to the large government

interventions in this period, these effects are also evident during the 2007-2009 recession as well. These findings are quantitatively significant. For example, our model produces forecast errors of up to 100 basis points in cumulative default rates over a two-year horizon during the Covid period, representing approximately a 25 percent overstatement relative to actual outcomes.

The second purpose of this paper is to apply discrete-time Markov transition methods to modeling auto default. This technique has been used in the mortgage lending literature, but less so in the auto lending literature. In auto lending, many loans become delinquent before going into default, but a significant fraction either exit for non-default reasons such as prepayment or return to being current after becoming delinquent. Both prepayment and default should be jointly accounted for to avoid potentially large errors in estimating transition probabilities (Deng et al., 2000). The Markov transition framework jointly estimates the transition probabilities between all the states and uses the information in these transition estimates to forecast defaults several periods ahead.

The final purpose of this paper is to add to the auto lending literature by demonstrating the importance of different lending channels for default, that is, whether a loan is made by a credit union, bank, or auto finance company. We document significant segmentation in auto lending by lender type with subprime lending concentrated in the auto finance sector. Furthermore, even when observables are accounted for, loan performance differs by lender type. For example, subprime loans originated by auto finance lenders are significantly riskier than subprime loans originated by other channels. The safest loans are those originated by credit unions and banks. We are aware of only two other papers that document heterogeneity in auto loan performance between banks, credit unions, and auto finance lenders.¹

¹A few papers in the literature have focused on lender effects, including An et al. (2020), Heitfield and Sabarwal (2004), and Katcher et al. (2024). However, these papers either did not identify the lender as a bank, credit union, or auto finance lender, or were focused on loan pricing rather than loan performance.

4.2 Background on the Auto Loan Market

The auto loan market is the second largest consumer credit market in the United States. In the fourth quarter of 2025 there was \$1.56 trillion of outstanding auto loans.²

Most auto loans are made through car dealers. A person buying a car goes to a dealer, and after the person decides to purchase a car, the dealer will present the buyer with some auto loan options. Dealers work with a variety of lenders. However, there are significant differences in auto loans by type of lender. The left panel of Figure 4.2 shows the percent of loans originated by lender types in our data, including auto finance companies, banks, and credit unions. Since 2000, banks' market share of the auto loan market has declined, while credit unions' market share has increased steadily. Prior work has documented that auto finance companies specializing in lending to the subprime segment of the market played an increasing role in financing that borrower segment during the pandemic (Canals-Cerda and Lee, 2021). The right panel of Figure 4.2 shows that auto finance companies originate the largest share of subprime auto loans.

Not only has the composition of lenders changed over time, but so has the distribution of auto loan origination credit risk. Figure 4.3 plots the Equifax Risk Score (henceforth "risk score") in our data from 2000-2025. Over that period, the aggregate share of loans in higher risk score groups has increased. Roughly 20 percent of loans were above a risk score of 760 in 2000, but this increased to roughly 40 percent of loans by 2025.

Historically, auto performance has followed the economy. Figure 4.4 shows the actual transition rates over 2000-2025. This period covers three recessions. The mild 2001 recession, the deep recession of 2007-2009, and the short Covid recession in 2020.³ Unlike in the other two recessions, auto defaults did not increase in 2020 and 2021; instead, default rates decreased. The decoupling of unemployment and auto loan risk outcomes in 2020 has been attributed to large government transfers and lender forbearance policies that appeared to

²Z.1 Financial Accounts of the United States.

³Transition rates for the first few periods in our sample are distorted due to censoring in the first few years of our dataset.

mitigate credit risks (for example, Bakshi, Tanya and Rose, Jonathan (2021); Dettling and Lambie-Hanson (2021)). Forbearance usage, in particular, rose to historically high levels in 2020 (Figure 4.5). These policies appeared to have distorted the traditional relationship between auto loan performance and unemployment. Some of this was due to forbearance, but as we will discuss, most of the better performance was not.

4.3 Literature

The literature on auto loan default probabilities is small, in part due to the lack of loan-level auto data that have been historically available to researchers. The earliest papers include Agarwal and Ambrose (2007), Agarwal (2008), and Heitfield and Sabarwal (2004), who model auto default and prepayment as a function of loan age, loan and borrower characteristics, and macroeconomic variables.⁴

Broadly, the literature on loan loss forecasting and stress testing find that unemployment and other macroeconomic variables are important for modeling consumer loan default.⁵ Most auto lending models include macroeconomic shocks. For example, Agarwal (2008) and Heitfield and Sabarwal (2004) found that default rates are sensitive to unemployment shocks, but prepayment rates are less so. Heitfield and Sabarwal (2004) also found that auto loan prepayment rates are insensitive to market interest rates. The Federal Reserve auto lending stress test includes unemployment and house price changes in its model.⁶

One challenge with modeling the effects of macroeconomic variables is the potential instability of estimates. Gross and Población (2019) and Guerrieri and Welch (2012) document that macroeconomic variables can lead to large amounts of uncertainty in reduced form statistical models of credit risk and bank performance, so should be used with caution. The

⁴Motivated by the recent increase in the length of auto loan terms, several recent studies find that loan age is important in explaining default risk, with loans of greater than five years having worse outcomes (Katcher et al. (2024), Guo et al. (2022); Wu et al. (2018)). We do not analyze this loan characteristic due to data limitations in our dataset.

⁵See, for example, Covas et al. (2014) and Hirtle et al. (2016).

⁶See <https://www.federalreserve.gov/supervisionreg/files/credit-risk-models.pdf>.

study of auto lending stress tests by Wu et al. (2018) shows significant differences in estimates of coefficients on macro variables in pre-GFC and post-GFC samples, though they use a different dataset for each period. They speculate that this is because they have only one recession in their sample. Unlike Wu et al. (2018) we use a single dataset and show that this result is not due to different data sources. Furthermore, we use rolling regressions that show how the estimate changes over time and why it changes. Finally, since our analysis covers two recessions, it shows that this instability is not unique to one recession.⁷

Much of the auto lending literature estimates discrete-time models that focus only on the exit states, such as conditional probability of default (PD) or competing risks models (An et al. (2020); Agarwal (2008); Heitfield and Sabarwal (2004); Guo et al. (2022); Wu et al. (2018)). The literature that applies Markov transition matrices to loan-level data typically does so on mortgage data (Douglas Smith and Lawrence (1995) and Smith et al. (1996), Chen et al. (2020)). One exception in the auto lending literature is Smith and Jin (2007). While that paper estimates Markov transition matrices, it does so on auto lease data and does not examine the effect of macroeconomic variables.

The last strand of the literature related to our work covers risk from different lender channels in auto lending. An et al. (2020) find that the lender channel is important in explaining default probabilities, with credit unions and banks facing relatively less risk of default when compared to auto finance companies. While nearly all studies establish a link between borrower credit quality and default probability, Heitfield and Sabarwal (2004) discuss the heterogeneous risks that exist for lenders who target different segments of the subprime market. Katcher et al. (2024) document differential prepayment rates and relationships between maturity and interest rates across lender types against a backdrop of increasing auto loan maturities over time. Our analysis builds on this analysis by further analyzing the effect of the lender channel.

⁷The earliest part of the sample also covers the 2000-2001 recession, but we start our rolling regressions in 2006 so that we have at least one recession in the first rolling estimation. This produces a single set of coefficients for the period covering 2000-2006, so we do not examine potential instability after the 2000-2001 recession.

4.4 Data

We use loan-level data on automobile loans from the FRBNY Consumer Credit Panel / Equifax Data (CCP) from June 2000-June 2025. This panel is an anonymized 5 percent national sample of all consumers who have a credit history reported to Equifax. The data contain auto loan and borrower characteristics over the life of the loan.

The data are semi-annual prior to 2017, quarterly over 2017-2019, and monthly beginning in 2020. We use semi-annual samples of the data for our analysis to be consistent throughout our sample period. The semi-annual samples in the CCP are as of June 1 and December 1 in each year. Furthermore, prior to 2015, the dataset only reports three broad types of lenders: auto finance companies, banks, and credit unions, so we use this classification of lender type.

We take several steps to clean the data prior to model estimation. First, the consumers the CCP follows are called primary borrowers. Some households include other people who also borrow. These other borrowers are not necessarily followed over time, so we drop them from our sample. Second, we drop loans from US territories, because these areas have few observations in the data. Third, we remove observations with missing or zero values for borrower age, Equifax Risk Score, or outstanding balances. Fourth, we remove leases, loans observed for greater than 9 years, loans where the borrower died during the life of the loan, loans with duplicated identifiers in a given time period, and a small number of loans with original balances of less than \$5,000 or more than \$100,000.⁸

We categorize a loan as being in one of four possible states. A loan is *current* if the payment is less than 30 days overdue. A loan is *delinquent* if the loan payment is 30 - 89 days late. A loan is in *default* if it is at least 90 days delinquent, repossessed, bankrupt, or charged off. Finally, a loan is a *non-default exit* if it exits and did not default. Non-default exit may include loans in which the balance is fully paid at term, is prepaid or refinanced, the borrower died, or the loan was sold by the lender. The reason for a non-default exit

⁸For example, the largest 0.001 percent of loans in the data carried balances greater than \$90,000.

is usually unobservable in our data.⁹ Furthermore, it is possible that some of these non-default exits are really defaults due to a loan going delinquent and exiting between periods of observation. To account for this possibility, we follow the procedure used by the Federal Reserve stress test, which is to make assumptions based on the size and status of the loan in the previous period along with the borrower’s credit score.¹⁰ Once a loan has exited the sample for default or non-default reasons, it exits our dataset.

We use borrower characteristics, loan features, and state-level macroeconomic variables in our estimation. The borrower characteristics are age, Equifax Risk Score, and whether the borrower has a mortgage. The loan features are the age of the loan, the outstanding loan balance, whether the loan is held jointly with someone else, and the type of lender. Outstanding loan balances are deflated by the consumer price index (CPI) using June 1999 as the base period. Other than the lender type, these variables are similar to the variables used in the literature (Agarwal (2008); An et al. (2020); Guo et al. (2022); Heitfield and Sabarwal (2004); Wu et al. (2018)). The loan and borrower age, mortgage status, and loan balance for each loan are allowed to vary over time for each loan in our model. We use the loan origination risk score of the borrower because movements in risk score may correlate with changes in other model variables, such as macroeconomic outcomes. We categorize the origination risk score into groups that broadly correspond to credit risk groupings typically used by lenders, spanning from super prime (an Equifax Risk Score greater than 760) to deep subprime (an Equifax Risk Score less than 560).¹¹ All other loan and borrower characteristics in the model are constant over the life of the loan.

The state-level macroeconomic variables we use are seasonally adjusted unemployment data from the Bureau of Labor Statistics (BLS) and the all-transactions house price index

⁹The CCP does not always report if a borrower dies. When it is reported, we drop the loan.

¹⁰For the details of the Federal Reserve’s process, see documentation from Board of Governors of the Federal Reserve System <https://www.federalreserve.gov/supervisionreg/files/credit-risk-models.pdf>.

¹¹Because the Equifax risk score is based on an internal model used by Equifax, we consider the possibility that the groupings we use may not match the credit score ranges commonly used. For robustness, we also categorize borrowers based on the origination quintiles of their Equifax Risk Score. Our results do not change using either approach.

(HPI) from the Federal Housing Finance Agency. We follow Wu et al. (2018) and measure these variables in terms of year-over-year changes. The Federal Reserve’s stress test also uses year-over-year changes in these two variables in its stress testing model, but their model also includes the level of unemployment as well as interactions of these variables with credit scores.¹² In our model, we find a negative coefficient on the level of unemployment associated with transitions to default, so we drop it from our model. Furthermore, given the number of other variables in our model, we do not interact unemployment change with credit score groups. We also exclude state fixed effects from our model because we find that these impair the accuracy of model predictions, and are likely highly correlated with the state-level macroeconomic variables. Finally, we tried several other potentially relevant macroeconomic variables such as Treasury rates, the prime rate, and used car price indices, but these were not statistically significant or meaningful in our estimation results, so we leave them out.

After cleaning and data construction, we have 296,388 unique loans and 1,668,736 total observations. Table 4.1 describes the variables used in our empirical model. Table 4.2 provides summary statistics for the transition rates and variables used in the regression analysis over the full sample from 2000-2025.

4.5 Empirical Framework

We use a discrete-time Markov transition framework where conditional probabilities of transition are estimated using a multinomial logit model. As discussed above, this approach has been used in the mortgage literature, but less so in the auto lending literature.

The benefit of the transition matrix framework is that it takes into account movement through the delinquency state and into the non-default exit state. Default modeling is often used to construct a measure of cumulative default over a longer horizon in order to assess the need for additional loss-absorbing capital. Multi-period-ahead forecasting using the transition matrix approach takes into account the probabilistic movements between states

¹²See <https://www.federalreserve.gov/supervisionreg/files/credit-risk-models.pdf>.

when constructing a cumulative default probability.

The Markov approach uses a recursive structure for forecasting the likelihood of transitioning between states. For each set of possible borrowers, lenders, and macroeconomic shocks let $\mathbf{x}_{it}$ be a vector of characteristics for loan i and let $\mathbf{z}_t$ be a vector of macroeconomic shocks, at time t . Let $\mathbf{y}_t$ be the vector of possible states for a loan. Conditioning the state transitions on $\mathbf{x}$ and $\mathbf{z}$, the Markov property assumes that

$$\begin{aligned} \text{prob}(\mathbf{y}_{it+1} = k | \mathbf{y}_{it} = j, \mathbf{y}_{it-1} = j_{t-1}, \dots, \mathbf{y}_{it-k} = j_{t-k}; \mathbf{x}_{it}, \mathbf{z}_t) \\ = \text{prob}(\mathbf{y}_{it+1} = k | \mathbf{y}_{it} = j; \mathbf{x}_{it}, \mathbf{z}_t) = p_{jk}(\mathbf{x}_{it}, \mathbf{z}_t). \end{aligned}$$

We let $\mathbf{P}(\mathbf{x}_{it}, \mathbf{z}_t)$ be the matrix of transition probabilities as a function of loan and borrower characteristics and macroeconomic shocks at time t .

In our analysis, there are four states. A loan can be current (c), delinquent (dq), default (df), or a non-default exit (e). The default and non-default exit states are absorbing states. The transition matrix $\mathbf{P}(\mathbf{x}_{it}, \mathbf{z}_t)$ contains elements $p_{jk}(\mathbf{x}_{it}, \mathbf{z}_t)$ that describe each of the possible transition probabilities at time t ,

$$\mathbf{P}(\mathbf{x}_{it}, \mathbf{z}_t) = \begin{pmatrix} p_{c,c}(\mathbf{x}_{it}, \mathbf{z}_t) & p_{c,dq}(\mathbf{x}_{it}, \mathbf{z}_t) & p_{c,df}(\mathbf{x}_{it}, \mathbf{z}_t) & p_{c,e}(\mathbf{x}_{it}, \mathbf{z}_t) \\ p_{dq,c}(\mathbf{x}_{it}, \mathbf{z}_t) & p_{dq,dq}(\mathbf{x}_{it}, \mathbf{z}_t) & p_{dq,df}(\mathbf{x}_{it}, \mathbf{z}_t) & p_{dq,e}(\mathbf{x}_{it}, \mathbf{z}_t) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4.1)$$

At this stage, the problem is to estimate the transition matrices $\mathbf{P}(\mathbf{x}_{it}, \mathbf{z}_t)$. Given the need to estimate probabilities across states and the fact that the data represent semiannual observations, we estimate conditional transition probabilities via maximum likelihood using a discrete time multinomial logit model. Pooling the observations, the likelihood function of

the multinomial logistic regression is

$$\begin{aligned}
 L &= \prod_{i=1}^N \prod_{t=1}^T P(y_{it} = j \mid \mathbf{x}_{it}, \mathbf{z}_t)^{\mathbb{1}(y_{it}=j)} \\
 &= \prod_{i=1}^N \prod_{t=1}^T \prod_{j=1}^4 \left[\frac{\exp(\mathbf{x}'_{it}\boldsymbol{\beta}_j + \mathbf{z}'_t\boldsymbol{\gamma}_j)}{\sum_{k=1}^4 \exp(\mathbf{x}'_{it}\boldsymbol{\beta}_k + \mathbf{z}'_t\boldsymbol{\gamma}_k)} \right]^{\mathbb{1}(y_{it}=j)}
 \end{aligned}$$

The model assumes that the conditional log odds of transitioning to state k depend linearly on a vector of loan and borrower characteristics $\mathbf{x}_{it}$ for loan i and a vector of macroeconomic variables $\mathbf{z}_t$ with corresponding coefficient vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. Let $p_{i,jk}(\mathbf{x}_{it}, \mathbf{z}_t)$ be the probability that loan i will transition from state j to state k at time t given $\mathbf{x}_{it}$ and $\mathbf{z}_t$. The model estimates

$$\ln\left(\frac{p_{i,jk}(\mathbf{x}_{it}, \mathbf{z}_t)}{p_{i,jj}(\mathbf{x}_{it}, \mathbf{z}_t)}\right) = \mathbf{x}_{it}\boldsymbol{\beta}_{jk} + \mathbf{z}_t\boldsymbol{\gamma}_{jk} \quad \text{for } k \neq j, \quad (4.2)$$

which gives

$$\frac{p_{i,jk}(\mathbf{x}_{it}, \mathbf{z}_t)}{p_{i,jj}(\mathbf{x}_{it}, \mathbf{z}_t)} = \exp(\mathbf{x}_{it}\boldsymbol{\beta}_{jk})\exp(\mathbf{z}_t\boldsymbol{\gamma}_{jk}) \quad \text{for } k \neq j$$

Exponentiation of the dependent variable is the probability of transition to state k relative to the probability of staying in the baseline state j . This *relative risk* (RR) is an intuitive way to interpret the coefficient estimates of the multinomial logit. A coefficient greater than 1 means that the relative probability of transitioning to a state increased more than the probability of staying in the same state, holding other covariates fixed. We define it as

$$RR_{i,k}(\mathbf{x}_{it}, \mathbf{z}_t) = \frac{p_{i,jk}(\mathbf{x}_{it}, \mathbf{z}_t)}{p_{i,jj}(\mathbf{x}_{it}, \mathbf{z}_t)} \quad \text{for } k \neq j.$$

The effect of a 1-unit change in x_{it} on RR is

$$\frac{RR_{i,k}(x_{it} + 1, \mathbf{z}_t)}{RR_{i,k}(x_{it}, \mathbf{z}_t)} = \frac{\exp(\beta_{ik}(x_{it} + 1))}{\exp(\beta_{ik}x_{it})} = \exp(\beta_{ik})$$

with a similar equation for the effect of a 1-unit change in z .

Finally, solving for the conditional transition probabilities that enter into $\mathbf{P}(t)$,

$$p_{i,jk}(\mathbf{x}_{it}, \mathbf{z}_t) = \frac{\exp(\mathbf{x}_{it}\boldsymbol{\beta}_{jk})\exp(\mathbf{z}_t\boldsymbol{\gamma}_{jk})}{1 + \sum_{k \neq j} \exp(\mathbf{x}_{i,t-1}\boldsymbol{\beta}_{jk})\exp(\mathbf{z}_t\boldsymbol{\gamma}_{jk})} \quad (4.3)$$

$$p_{i,jj}(\mathbf{x}_{it}, \mathbf{z}_t) = \frac{1}{1 + \sum_{k \neq j} \exp(\mathbf{x}_{it}\boldsymbol{\beta}_{jk})\exp(\mathbf{z}_t\boldsymbol{\gamma}_{jk})} \quad (4.4)$$

The multinomial logit model jointly estimates parameters of each transition equation such that the estimated probabilities are constrained to be between 0 and 1 and sum to 1. Because we are modeling the entire state space, we estimate a separate multinomial logistic regression for the two starting states, current and delinquent. The estimated probabilities correspond to the first two rows of (4.1). In all estimations, we cluster standard errors at the loan level.

Our multinomial logit framework estimates discrete-time hazard functions for transitions to the absorbing states of default and exit. In the discrete-time hazard literature, the hazard is defined as the conditional probability of experiencing an event in period t , given survival to period t . In our competing risks framework with multiple transient states (current and delinquent), the cause-specific hazard functions are simply the transition probabilities to the absorbing states:

$$h_{i,j \rightarrow k}(t|\mathbf{x}_{it}, \mathbf{z}_t) = p_{i,jk}(\mathbf{x}_{it}, \mathbf{z}_t) \quad \text{for } j \in \{c, dq\}, \quad k \in \{df, e\}.$$

The baseline hazard is captured through loan age, which enters the linear predictor in equation (4.2). We use a quadratic specification of loan age (in months) to flexibly model

duration dependence. As Lancaster (1990) demonstrates, misspecification of the baseline hazard can lead to biased parameter estimates. We compared several functional forms and found the quadratic specification provided the best fit in terms of the root mean squared error, McFadden’s pseudo- R^2 , and the Bayesian information criteria (see Appendix 4.12.1).

4.5.1 Multi-Period-Ahead Forecasts of Default

Estimating a cumulative probability of default or non-default exit over multiple periods is commonly used by lenders for loss forecasting and by regulators for stress testing. This exercise is straightforward in the Markov transition matrix framework, but requires assumptions on the data given that the probabilities are conditional on changing loan and macroeconomic variables. One method is to assume a path of macroeconomic variables while holding other non-deterministic variables constant. In this setting, lenders and regulators can assess the sensitivity of existing loan portfolios to macroeconomic shocks.

To estimate a cumulative probability of default and non-default exit, we need the initial distribution of states that loans are in, π_0 ; the initial distribution of loan and borrower characteristics, $(\mathbf{x}_{i0}, \mathbf{z}_0)$; how the deterministic characteristics change over time, e.g., loans age deterministically, which we call $\tilde{\mathbf{x}}_{i1}$ to reflect that only certain characteristics change over time; the time path of the macro shocks $\tilde{\mathbf{z}}_1$; and the transition matrix P . The h -period-ahead cumulative forecast of the distribution of states, π_{0+h} is calculated by

$$\pi_{0+h} = \pi_0 \mathbf{P}(\mathbf{x}_{i0}, \mathbf{z}_0) \mathbf{P}(\tilde{\mathbf{x}}_{i1}, \tilde{\mathbf{z}}_1) \cdots \mathbf{P}(\tilde{\mathbf{x}}_{ih}, \tilde{\mathbf{z}}_h) \quad (4.5)$$

4.6 Model Estimation

Using our chosen hazard specification, we first estimate the full model specification from 2000-2019 in order to understand relationships between model covariates and transitions prior to potential distortions from the pandemic period. Tables 4.3 and 4.4 show the results

of our estimation. In each table the coefficients represent the marginal effects on the relative risk of transition to a particular state compared with staying in the current state. For the categorical variables that we include in the model that have more than two levels, credit unions and origination risk scores greater than 760 are the reference groups for lender type and origination risk score, respectively.

4.6.1 Transitions from the Current State

We first focus on the estimates of the transition from the current state. The loan and borrower characteristics included in the model are almost all statistically significant at the 1 percent level. Borrowers who are younger, in lower risk score groups, non-mortgage holders, further from origination, and do not have a joint account have higher relative risks of transition into delinquency or default relative to remaining current. The impacts of low risk scores are particularly pronounced. For example, the odds of a deep subprime borrower transitioning from current to delinquency (default) relative to a transition to current for borrowers with a risk score of at least 760 increase by a factor of 7000 percent (2200 percent). The lender channel also shows meaningful effects, particularly for loans originated by auto finance companies. The odds of an auto finance loan transitioning to delinquency (default) relative to staying current and originating from a credit union are 70 percent (55 percent) higher.

Borrowers with the highest origination risk scores, mortgage holders, and borrowers with lower balances have relatively higher odds of transitioning to non-default exit. The absolute magnitude of estimated effects for risk score and lender type on the relative risk of transitioning to the non-default exit state is smaller than that those for transitioning to default. The odds of a deep subprime (<560) borrower transitioning to exit relative to a transition to current for borrowers with a risk score of at least 760 decrease by 17 percent. The relative risks of exit for loans that originate from auto finance lenders are 3 percent lower relative to staying current and originating from a credit union.

Unlike the coefficients on risk score and lender type, the coefficients on macroeconomic variables are not always statistically significant. A 1 percentage point increase in the year-over-year change in the unemployment rate is associated with a 9 percent increase in the relative risk of transitioning to default, but the other states do not show statistical significance. A 1 percentage point increase in the annual HPI slightly decreases the relative risk of transitioning to delinquency, and slightly increases the relative risk of transitioning to non-default exit. While these effects seem small compared to credit score or lender effects, large macroeconomic shocks during economic downturns can have larger effects on the relative risks. Indeed, this is the primary role for these variable in stress testing models that condition on tail macroeconomic outcomes.

4.6.2 Transitions from the Delinquent State

For transitions from the delinquent state, estimated relative risk coefficients on the borrower and loan characteristics are again generally statistically significant and intuitive. The patterns in the current state model are similar to those in the delinquent state model, but borrower age no longer matters statistically or in a meaningful way. Also, unlike starting in the current state, an increase in the loan age results in a decrease in the relative risk of transitioning into default for borrowers in the delinquent state. A loan originated from auto finance lenders increases the relative risks of transitioning to delinquency and default by 9 and 10 percent, respectively, relative to transitions into current by borrowers from credit unions. The relative risks of transitions into delinquency and default for loans originated with a risk score of less than 560 increase by 284 and 254 percent, respectively, relative to transitions into current by borrowers with an origination credit score higher than 760. For transitions into non-default exit, the results reverse such that the relative risks of exit are decreased for loans originated from auto finance companies and for borrowers with lower credit risk scores.

The estimated effects of macroeconomic variables are not always significant or intuitive.

For the year-over-year change in unemployment, the effect is only statistically significant for transitions to the exit state, but the effect suggests that an increase in unemployment increases the relative risks of non-default exit. As discussed earlier, it is possible that our data-cleaning procedure for identifying defaults from transitions to non-default exit misses some defaults. If that is the case, this measurement error could account for this finding. The estimated effects of a 1 percentage point increase in the annual HPI is statistically significant for the exit and default states, and suggests a slight increase in the relative risk of transition to exit and a slight decrease in the relative risk of transition to default.

4.7 Rolling Regressions, Predicted Transition Probabilities, and Coefficient Instability

We now run rolling regressions to estimate period-by-period aggregate transition probabilities and forecasted losses and then compare them to actual outcomes. We also examine how estimated coefficients and forecasted losses change as new data arrive and the model is re-estimated with the new data.

We use the period from 2000-2006 as the initial in-sample period. In our rolling predictions, we predict time $t + 1$ outcomes on the parameter estimates estimated on all data up to time t . We then re-estimate the model parameters through all data as of $t + 1$ and predict outcomes at $t + 2$, and so on. This process continues until the last data point in our sample, ending in 2025.

Figure 4.6 plots the pseudo out-of-sample one-period-ahead forecasts (henceforth out-of-sample) compared with actual outcomes. The rows of the plots correspond to the first two rows of equation (4.1), summed over all loans at each time t . The red line in the first row plots the transition probability for current loans, given the characteristics of current loans at time t , and the red line in the second row plots the transition probability for delinquent loans, given the characteristics of delinquent loans at time t . Red shaded areas represent 95

percent confidence intervals around each prediction. The solid black line is the fraction of current or delinquent loans that actually transitioned to each new state from time t to time $t + 1$. The dashed vertical line is December 2006, which is where the out-of-sample forecasts begin. (Forecasts in the plots prior to December 2006 are in-sample forecasts.)

The multinomial logit generally captures the patterns in the actual transition frequencies during the in-sample period and the out-of-sample rolling forecasts. After 2006, the model slightly underpredicts the transition from current to current in several periods, while it overpredicts the transitions from current to non-default exit. Similarly, the model slightly underpredicts transitions from delinquent to default, while it overpredicts from delinquent to exit.

The most notable forecast errors occur in the first semiannual period of 2020 for transitions from current to default and delinquent to exit. In this period, actual default rates declined, while the model predicted sharp increases in the transition probabilities. The next largest forecast errors occur during the GFC. Unlike during the GFC, the default and exit outcomes generally move in the same direction as the model predictions despite the model overpredicting the outcomes.

Forecast errors can arise when conditional relationships between the transitions and covariates fundamentally change. Given the size and scope of the federal government's response to Covid, this is a period where one would expect changes in these relationships. Our rolling coefficients allow us to examine how large these effects are on the estimated relationships. We focus on the coefficients in the current to default transition because current loans are a much larger share of loans than delinquent loans and for one-period-ahead forecasts it will have the largest impact on defaults.

We first examine the rolling standard deviation of the semiannual change in standardized rolling relative risk coefficients to look for evidence of coefficient volatility. Figure 4.7 shows the rolling standard deviations for transitions from current to default. With the exception of the model intercept, all of the loan and borrower covariates typically have a semiannual

change of less than 1 standard deviation in 2020. The highest volatility in the first semi-annual period of 2020 is in the seasonal fixed effect; presumably this rolling seasonal effect is picking up some of the pandemic-related changes during this period. The next highest volatility is in the unemployment coefficient, which changes by 1.3 standard deviations in June 2020 and the HPI change coefficient is similar. The volatility of both macro coefficients is usually small except during the GFC and Covid shocks. Figure 4.8 shows the rolling standard deviations for transitions from delinquent to exit. The coefficient instability for unemployment (nearly 3 standard deviations) was larger than for HPI.

While the coefficients for unemployment and HPI show volatility spikes in June 2020, we focus on unemployment for two reasons. First, unemployment experienced a much larger shock in 2020 (Figure 4.9) compared to the relatively stable HPI (Figure 4.10), which did not increase sharply until 2021. Second, the volatility of the unemployment coefficient is larger than that of the HPI coefficient when looking across both the GFC and Covid recessions, suggesting a more substantial change in default risk in response to unemployment shocks. This suggests unemployment as the primary driver of forecast errors during the two recessions in the rolling results.

Figure 4.11 reports the estimates of the unemployment coefficients through time in the rolling regressions. There is a large decline in this coefficient for the current-exit and current-default transitions during the GFC, and another sharp decline across four of the transition probabilities in 2020. (This drop explains the increase in volatility of the unemployment coefficient in Figures 4.7 and 4.8.) Given the large forecast errors associated with transitions from current-default in those two periods, we focus on the coefficient associated with that transition. At the beginning of the rolling estimation in December 2006, the marginal effect of a 1 percentage point increase in unemployment was associated with an increase in the relative risks of roughly 16 percent. After the GFC and before the pandemic, this effect drops to slightly less than 10 percent, which then declines again to about 3 percent in the post-pandemic period. Furthermore, the level of the effect after each decline is persistent, es-

pecially in the post-2020 period. After 2020, the coefficient stays low in the rolling regression through the end of our sample in 2025.

It is instructive to understand how these coefficients may have impacted the riskiest segments of the borrower population. We compute the marginal effects for a representative subprime borrower across a grid of values for annual unemployment changes to further illustrate the impact of an unemployment shock. We define a representative subprime borrower where the loan is originated by a bank and has an origination credit score of less than 560 while all other variables are held at their respective mean value across all borrowers. Figure 4.12 plots the increase in probability of transition from current to default for given year-over-year changes in US state unemployment. The x-axis represents a range of potential year-over-year changes in US state unemployment, the y-axis represents the marginal effect of a given change on the probability of transition from current to default.

In 2006, positive changes in unemployment of 10 percentage points (roughly the magnitude of the average unemployment change during the GFC) were associated with roughly a 1 percent increase in the probability of default. The size of the effect continues to increase as unemployment increases. Comparing this to the marginal effects generated across the same grid of annual unemployment increases in 2019 and 2021, it is only in 2006 that unemployment has a meaningful effect on default probabilities at relatively large shocks to the level of unemployment. There is a very small positive effect in 2019 between unemployment changes between 10 to 20 percent, but the effect is small relative to 2006. By 2021, the effect is effectively flat across the potential changes in annual unemployment.

These figures provide evidence of instability, namely, a decline in the relationship between unemployment and transitions to default after a large shock. The drop after 2020 should not be surprising. During Covid, there were two significant changes in the auto lending environment. First, there was a large spike in the use of forbearance by auto lenders. Second, the federal government made enormous transfers to households to offset the macroeconomic effects of Covid. These actions prevented many defaults, breaking down the traditional

relationship between unemployment and transitions to default. Nevertheless, as we will see, these two factors do not change the underlying finding about instability.

4.7.1 Forbearance

Forbearance agreements between a borrower and lender allow for temporary reduction or suspension of loan repayment, often requested by borrowers during periods of economic hardships. One concern with our errors in forecasting default in 2020 is that they may be driven by borrowers who entered into forbearance, so that the borrowers would have defaulted had the contractual obligations of their loan not been changed by lenders. If characteristics of the borrower or loan are correlated with forbearance usage, or if forbearance usage occurred disproportionately in US states with larger spikes in unemployment, then some portion of our 2020 forecast errors could be a result of forbearance. As previously shown, forbearance increased to historically high levels in the same period of our default forecast errors.

To assess the degree to which loans that entered into forbearance change the forecast results, we compare the out-of-sample predicted probabilities between the full set of loans in our data and the set of loans that never enter into forbearance. Figure 4.13 shows the results. The red line indicates the average predictions from the rolling regressions on the full sample of loans; the blue line shows the average predictions for non-forbearance loans. The two forecasts almost perfectly overlap, including the period of the large current-default and delinquent-exit forecast errors in 2020. This suggests that forbearance does not explain our findings.

4.7.2 Controlling for 2020 Interventions

As we discussed, federal government interventions were enormous during Covid and far exceeded any other peacetime period in US history, which should alter the relationship between unemployment increases and defaults. To the extent that the pandemic represents an idiosyncratic period not representative of other recessionary periods, a simple way to control

for this period is to include a dummy variable in the model for the pandemic period. We re-estimate (4.2) for the current and delinquent starting states, but now add a dummy variable equal to 1 during the semiannual periods from 2020 through 2021. If the dummy variable absorbs pandemic-specific factors that temporarily altered the unemployment-transition relationship, then the unemployment coefficients should better reflect the stable structural relationship that holds in normal periods.

Figure 4.14 shows the rolling coefficient estimates on the unemployment variable in the model that includes dummy variables for the pandemic period. We still find evidence of a declining relationship between unemployment and transitions to default. Relative to our prior results, the unemployment coefficient for the current-to-default transition declines more slowly over the 2020-2021 period and settles at a slightly higher level, but qualitatively the effect remains. We take this as evidence that the effects of the pandemic intervention are persistent on auto lending default performance, so a pandemic period dummy does not fully account for these effects.

4.8 Cumulative Forecasts

Lenders and regulators often construct cumulative forecasts of payoffs and default over time in order to project future revenues and the amount of loss-absorbing funds needed over a longer horizon. A two-year cumulative default rate is typically used in stress testing applications. For these types of exercises, our Markov structure is valuable. We assess the impact of the structural break in the relationship between unemployment and transitions to absorbing states on two-year cumulative forecasts. As we showed in Figure 4.6, forecasted transitions from current to delinquent were much higher than actual transitions. In this section, we break down the effect on forecasts from the structural break in the unemployment coefficient versus other coefficients.

In the predictions, we assume that the future path of macroeconomic variables is known,

so that we are able to directly compare the model predictions to actual outcomes. For this exercise, additional assumptions are needed around the future path of the remaining model covariates. For borrower and loan characteristics that are deterministic (such as loan age and borrower age) we allow them to vary as time progresses in the forecast period, but non-deterministic variables that vary over time (such as outstanding loan balance and mortgage status) are held constant as of the latest observed data. Using these assumptions on the path of future data, (4.3) can be used to construct the time-varying predicted transition matrices and (4.5) can be used to construct a cumulative forecast based on those matrices.

We decompose the impact of the changing relationships between unemployment and transitions to default and exit by varying the parameters used to produce the cumulative forecasts. We use three sets of model parameters that we refer to as pre-date, post-date, and hybrid. The pre-date model coefficients are estimated as of December 2019. The post-date model uses the estimated coefficients as of June 2020. The hybrid model uses all of the post-date model parameters except for unemployment, and instead uses the unemployment coefficient as of December 2019. A comparison of these cumulative forecasts allows us to estimate how much of the variation in cumulative forecasts comes from revised relationships between unemployment versus the other model parameters.

Figure 4.15 shows the results of the default (left panel) and exit (right panel) forecasts as they accumulate from the first forecast period of June 2020 through December 2021, and Figure 4.16 plots the differences in the forecasts relative to actual outcomes at the end of the two-year period. The default predictions for the pre-date model simulates what was possible to forecast as of December 2019 if the path of unemployment and the HPI were known. The divergence between the pre-date model and actual outcomes by the end of the two-year cumulative forecast is roughly 100 basis points. As shown by Figure 4.16, this difference is economically significant and represents roughly a 25 percent overstatement in the predicted cumulative default relative to actual outcomes.

The predictions from the post-date model still produce a cumulative forecast higher than

the actual outcome, but the differential is reduced by roughly two-thirds. Differences between the pre-date and post-date models point to the changes in coefficients as explaining much of the difference. The hybrid model breaks down the effect of changes in different coefficients. The hybrid model predictions are roughly halfway between the pre-date and post-date model predictions, which suggests that the change in the unemployment coefficient accounts for half of the difference in predictions and the other coefficients account for the other half.

4.8.1 Borrower Risk Score and Lender Channels

Our model controls for borrower risk score and origination lender type, so we use that to further examine how the cumulative forecasts differ across these sources of loan heterogeneity. Results along these dimensions will be correlated because banks and credit unions generally originate loans to higher quality borrowers relative to auto finance companies. Figure 4.17 plots default rates in our data by the origination risk score categories we use, ranging from the least risky (760+) to the riskiest deep subprime borrowers (< 560). While default rates are generally near zero for the least risky borrowers, they often increase more than 5 percentage points for the riskiest borrowers. Figure 4.18 plots default rates for loans originated by the lender categories that we observe in our data. Auto finance loans typically have default rates above 1.5 percent, and their default rates are usually at least a full percentage point higher than those of banks or credit unions.

To highlight where the default forecast errors are largest using the model coefficients as of December 2019, we examine forecast errors in both absolute differences and as a percent of the actual default rate. Figure 4.19 plots the cumulative default transition forecast differences relative to actual outcomes by original risk score groupings. In absolute terms, the forecast errors are largest for the riskiest borrowers, but in percent terms the forecast errors are largest for the least risky borrowers. The absolute difference is reflective of the higher default rates for low-quality borrowers. Small errors, however, can be large in percentage terms for the highest quality borrowers because they have very low default rates.

Figure 4.20 plots the same analysis of cumulative forecast differences for origination lender types. Similar to the risk score results, the absolute errors are largest for auto finance lenders, which originate loans to a larger share of risky borrowers. The percentage difference is relatively higher for banks and credit unions, which hold less risk in their auto loan portfolios and face lower default rates.

Stress testing is concerned with forecasted losses in levels so inaccuracies in level forecasts can be economically meaningful for lenders or regulators in terms of capital levels. Alternatively, percentage forecast errors relative to realized default rates highlight model calibration issues in the statistical sense, that is, whether predicted probabilities match actual default frequencies across the risk distribution. For example, because default rates among the safest borrowers are low, small deviations can lead to large percentage errors, mischaracterizing the macroeconomic sensitivity of safer borrowers and distorting risk differentiation across the credit risk distribution. The percentage forecast errors we find are notable because the majority of auto lending is made to higher quality borrowers, so that small absolute errors can aggregate to large errors for an overall portfolio of loans. This can have large economic implications for loss planning for lenders in addition to the large absolute errors for the riskiest segments of the loan population.

4.8.2 Oaxaca-Blinder Decomposition

While we find evidence that the statistical relationship between unemployment and default changed in 2020 and that this change explains the large forecast error, it is possible that the forecast error is also due to changes in borrower and loan characteristics. To evaluate this possibility, we follow Fairlie (2005) and use a nonlinear Oaxaca-Blinder decomposition on our data and rolling coefficient estimates. This procedure involves computing average predicted probabilities from the data and model coefficients in two different periods, a baseline period and an alternative period. The total difference in the average predicted probabilities over the two periods is then decomposed into components explained by the changes in loan data,

macroeconomic data, and the estimated coefficients.

We set December 2019 as the baseline period and December 2021 as the alternative period. The predictions use various combinations of the borrower and loan characteristics $\mathbf{x}$, macroeconomic variables $\mathbf{z}$ and the estimated borrower and loan coefficients $\hat{\beta}$ and the estimated macro shock coefficients $\hat{\gamma}$ from each of the two time periods. Taking our model's average predicted probabilities for transitions from state j to state k as $\bar{p}_{jk}$, define the total change in average predicted probabilities between the baseline 2019 period b and alternative 2021 period a

$$\Delta^{Total} = \bar{p}_{jk}(\mathbf{x}_a, \mathbf{z}_a; \hat{\beta}_a, \hat{\gamma}_a) - \bar{p}_{jk}(\mathbf{x}_b, \mathbf{z}_b; \hat{\beta}_b, \hat{\gamma}_b) \quad (4.6)$$

$$\Delta^{Borrower} = \bar{p}_{jk}(\mathbf{x}_a, \mathbf{z}_b; \hat{\beta}_a, \hat{\gamma}_a) - \bar{p}_{jk}(\mathbf{x}_b, \mathbf{z}_b; \hat{\beta}_a, \hat{\gamma}_a) \quad (4.7)$$

$$\Delta^{Macro} = \bar{p}_{jk}(\mathbf{x}_a, \mathbf{z}_a; \hat{\beta}_a, \hat{\gamma}_a) - \bar{p}_{jk}(\mathbf{x}_a, \mathbf{z}_b; \hat{\beta}_a, \hat{\gamma}_a) \quad (4.8)$$

$$\Delta^{Coefficient} = \bar{p}_{jk}(\mathbf{x}_b, \mathbf{z}_b; \hat{\beta}_a, \hat{\gamma}_a) - \bar{p}_{jk}(\mathbf{x}_b, \mathbf{z}_b; \hat{\beta}_b, \hat{\gamma}_b) \quad (4.9)$$

This specification examines what explains the difference between predictions for 2021 made in 2019 and the in-sample prediction for 2021 made using the rolling regression estimated through 2021. The coefficient effect measures how much the model change matters when applied to 2019 conditions, while the borrower and macro effects measure how much data changes matter when evaluated with the new model. In this decomposition, coefficients are updated holding the data fixed at the baseline. Then, holding coefficients fixed at the alternative levels, borrower composition is updated holding macro variables fixed at baseline levels and macro variables are updated holding borrower characteristics fixed at alternative levels. This allows the total predicted change to be decomposed as the sum of the components,

$$\Delta^{Total} = \Delta^{Borrower} + \Delta^{Macro} + \Delta^{Coefficient} \quad (4.10)$$

To examine the relative contributions to the total effects, Figure 4.21 plots the percent of the total effect explained by each component for the two absorbing states. The results show that macroeconomic changes dominate (60-70 percent), reflecting the massive unemployment spike from 3 to 13 percent. Even with the weakened unemployment sensitivity captured by 2021 coefficients, this shock is so large it overwhelms other factors. The coefficient effect (20-30 percent) captures the model change, but is smaller because it is evaluated on 2019 data when unemployment is low. The borrower composition effect is much smaller than the macroeconomic effect (18 percent), ruling out composition changes as the primary driver of the forecast errors.¹³

The decomposition results are consistent with a large macroeconomic shock that generates a large forecast error rather than the error being meaningfully impacted by changes in borrower composition. Taken together, our results show that the resulting forecast errors are generated by a large change in the macroeconomic data rather than a large change in the underlying population of borrowers.

4.8.3 Comparison to Binomial Discrete-Time Hazard Model

A simpler approach to estimating probabilities of default for discrete-time survival analysis or hazard modeling is to use a single binary outcome regression. Indeed, papers in the academic literature as well as regulatory stress testing models rely on this approach rather than constructing an entire transition matrix. The binomial discrete-time hazard model estimates the probability of default without taking into account probabilistic movement throughout the state space and the competing risk of non-default exit. The binomial approach may be appropriate when loan dynamics are relatively simple—for instance, when exit rates are low, cure rates are minimal, and state transitions are rare. However, when these simplifying assumptions are violated, the binomial model can produce biased forecasts.

Our historical transition rates across states motivate why we might care about the tran-

¹³Given that the effects are nonlinear, Fairlie (2005) notes that the ordering of the effects matters. An alternative ordering of the effects produces nearly identical results.

sition approach. On average, there is a 40 percent probability that loans in our sample that start as delinquent will “cure” to current, and more than a 12 percent probability that loans starting as current will exit without default. Failing to take these probabilities into account when estimating a cumulative probability can impact the accuracy of the estimates, especially in multi-period-ahead forecasts.

For comparison, we estimate a binomial discrete-time hazard model. The binomial model takes the form

$$\ln \left(\frac{p_{i,df}(t)}{1 - p_{i,df}(t)} \right) = \mathbf{x}_{it}\boldsymbol{\beta} + \mathbf{z}_t\boldsymbol{\gamma} \quad (4.11)$$

where $p_{i,df}(t)$ represents the 1-period-ahead probability of default for loan i at time t . Solving for $p_{i,df}$, the h -period-ahead cumulative probability of default P_{it}^h , accounting for the probability of surviving each previous period, can be recursively constructed as

$$\begin{aligned} P_{it}^1 &= p_{i,df}(t) \\ P_{it}^2 &= P_{it}^1 + (1 - P_{it}^1) \cdot p_{i,df}(t + 1) \\ &\vdots \\ P_{it}^h &= P_{it}^{h-1} + (1 - P_{it}^{h-1}) \cdot p_{i,df}(t + h - 1) \end{aligned}$$

Note that this specification accounts for survival (non-default) but not for exits or state transitions. A loan that exits in period $t = 2$ continues to enter the survival probability calculation in subsequent periods, and a delinquent loan that cures continues to face delinquent-state default probabilities. This contrasts with the transition matrix approach, where the state vector $\mathbf{P}(\mathbf{x}_{it}, \mathbf{z}_t)$ dynamically adjusts to reflect exits, cures, and risk transitions.

We estimate (4.11) using our data from 2000-2019 and use it to produce two-year cumulative forecasts over the period 2020-2021. As previously shown in Figure 4.4, actual

transitions to default declined during that period, but transitions from current to exit and from delinquent to current both increased. Figure 4.22 plots the path of the cumulative default predictions implied by the binomial logit model and the Markov transition and the actual cumulative default outcomes over the 2020-2021 period. The initial forecast in June 2020 is roughly equivalent between the two models. However, by the end of the forecast period, the binomial model produces a cumulative forecast error that is 34 percent higher than the actual cumulative outcome while the Markov transition matrix is 24 percent higher.

While the Markov model produces lower cumulative forecast errors during the 2020-2021 period, it is instructive to better understand if the Markov transition model may be more accurate in capturing the dynamics of borrower transitions across other periods in our sample. We use both models to produce a two-year cumulative forecast at each semiannual period in our sample, allowing the macroeconomic variables to evolve as they actually did and treating the loan characteristics as previously described. Table 4.5 shows the total cumulative forecast error comparison between the Markov transition matrix and the binomial model across all time periods. The Markov transition model has lower cumulative forecast errors across standard measures that include the mean absolute error (MAE), the mean absolute percentage error (MAPE), and root mean squared error (RMSE). Figure 4.23 plots the rolling two-year cumulative default forecast errors for the Markov transition matrix and binomial logit model at each date in our sample. The Markov transition model produces lower cumulative forecast errors in almost all periods, with the exception of several years from 2015-2017 when the exit transition forecasts deteriorated and in 2025.

These results imply that the Markov transition model's estimated overprediction of cumulative default rates using estimated model parameters as of 2019 are conservative relative to simpler modeling approaches that may be used by lenders and regulators. It also implies that the forecast errors in the binomial model will be driven not only by the unstable sensitivity of default rates to unemployment, but also by the fact that the binomial model does not account for all of the dynamic transitions across the state space.

4.9 Conclusion

We estimate a Markov transition model of auto loan performance over the period from 2000-2025 using a panel of semiannual loan-level data. The pandemic period was characterized by large changes in macroeconomic variables typically used to model loan performance under stress, but loan performance did not deteriorate in 2020-2021 in part due to large government transfers that offset the shocks. We examined the implications for loss forecasting using our model to produce forecasts in a rolling pseudo out-of-sample framework. We also document a lender channel of loan performance risk, where the riskiest loans are originated by the auto finance industry.

Using the Markov transition framework, we produce multi-period-ahead forecasts that take into account the probabilistic movements between states of loan performance. We find large single and multi-period-ahead default forecast errors during the pandemic period and instability in the estimated parameters on unemployment. The forecast errors are attributable the instability of the unemployment coefficient in the model that occurs after large aggregate shocks, first during the 2007-2009 recession and again during the 2020 Covid recession. Moreover we showed that the instability leads to persistently lower levels of sensitivity to unemployment shocks for loans that transition to default rather than eventually returning to prior levels. Counterfactual forecasts and Oaxaca-Blinder decompositions showed that the forecast errors and coefficient instability were primarily driven by the model learning the revised relationships between unemployment and loan performance outcomes. We show that forbearance does not appear to explain the unstable relationship during Covid, and robustness tests that attempt to control for the 2020-2021 period using dummy variables do not improve the forecast errors or parameter instability.

The results suggest that the relationship between unemployment and auto loan default is declining to a point where it is nearly meaningless for loan default predictions. If our model were used to evaluate policies around auto loan defaults, there are obvious connections to the Lucas critique (Lucas (1976)) given the reduced-form relationships used to model the

impact of macroeconomic outcomes on aggregate loan performance. Given how common unemployment is in reduced-form models of consumer loan loss forecasting, these findings are concerning for lenders and regulators that rely on these estimated relationships to correctly forecast future losses. Future work could examine how structural approaches to loan loss modeling may improve on forecasts of loan performance in periods where government interventions or changes in lending practices due to a recession change statistical relationships.

Bibliography

- Abdymomunov, Azamat, Jeffery Gerlach, and Yuji Sakurai (2023). “Interest Rate Risk in the U.S. Banking Sector”. en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: [10.2139/ssrn.4395529](https://doi.org/10.2139/ssrn.4395529). URL: <https://www.ssrn.com/abstract=4395529> (visited on 03/25/2024).
- Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone (Apr. 2019). “Vulnerable Growth”. en. In: *American Economic Review* 109.4, pp. 1263–1289. ISSN: 0002-8282. DOI: [10.1257/aer.20161923](https://doi.org/10.1257/aer.20161923). URL: <https://pubs.aeaweb.org/doi/10.1257/aer.20161923> (visited on 09/01/2023).
- Agarwal, Sumit (2008). “Determinants of Automobile Loan Default and Prepayment;” en. In: *Economic Perspectives* 32.3.
- Agarwal, Sumit and Brent W. Ambrose, eds. (2007). *Household Credit Usage*. en. New York: Palgrave Macmillan US. ISBN: 978-1-349-53946-8 978-0-230-60891-7. DOI: [10.1057/9780230608917](https://doi.org/10.1057/9780230608917). URL: <http://link.springer.com/10.1057/9780230608917> (visited on 02/17/2026).
- Aghamohammadi, A. and S. Mohammadi (Dec. 2017). “Bayesian analysis of penalized quantile regression for longitudinal data”. en. In: *Statistical Papers* 58.4, pp. 1035–1053. ISSN: 0932-5026, 1613-9798. DOI: [10.1007/s00362-015-0737-4](https://doi.org/10.1007/s00362-015-0737-4). URL: <http://link.springer.com/10.1007/s00362-015-0737-4> (visited on 03/25/2024).
- Alessandri, Piergiorgio and Benjamin D. Nelson (Feb. 2015). “Simple Banking: Profitability and the Yield Curve”. en. In: *Journal of Money, Credit and Banking* 47.1, pp. 143–175.

- ISSN: 00222879. DOI: [10.1111/jmcb.12172](https://doi.org/10.1111/jmcb.12172). URL: <https://onlinelibrary.wiley.com/doi/10.1111/jmcb.12172> (visited on 09/01/2023).
- Alhamzawi, Rahim, Keming Yu, and Dries F Benoit (2012). “Bayesian adaptive Lasso quantile regression”. In: *Statistical Modelling* 12.3, pp. 279–297. DOI: [10.1177/1471082X1101200304](https://doi.org/10.1177/1471082X1101200304).
- An, Xudong, Larry Cordell, and Sharon Tang (May 2020). *Extended Loan Terms and Auto Loan Default Risk*. en. Working paper (Federal Reserve Bank of Philadelphia) 20-18. Series: Working paper (Federal Reserve Bank of Philadelphia). Federal Reserve Bank of Philadelphia, pp. 20–18. DOI: [10.21799/frbp.wp.2020.18](https://doi.org/10.21799/frbp.wp.2020.18). URL: <https://www.philadelphiafed.org/-/media/frbp/assets/working-papers/2020/wp20-18.pdf> (visited on 02/16/2026).
- Angbazo, Lazarus (Jan. 1997). “Commercial bank net interest margins, default risk, interest-rate risk, and off-balance sheet banking”. en. In: *Journal of Banking & Finance* 21.1, pp. 55–87. ISSN: 03784266. DOI: [10.1016/S0378-4266\(96\)00025-8](https://doi.org/10.1016/S0378-4266(96)00025-8). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0378426696000258> (visited on 09/01/2023).
- Azzalini, Adelchi and Antonella Capitanio (May 2003). “Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t -Distribution”. en. In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 65.2, pp. 367–389. ISSN: 1369-7412, 1467-9868. DOI: [10.1111/1467-9868.00391](https://doi.org/10.1111/1467-9868.00391). URL: <https://academic.oup.com/jrsssb/article/65/2/367/7092856> (visited on 09/01/2023).
- Bakshi, Tanya and Rose, Jonathan (2021). “What happened to subprime auto loans during the COVID-19 pandemic”. In: *Chicago Fed Insights*.
- Begenau, Juliane and Erik Stafford (2022). “Unstable Inference from Banks’ Stable Net Interest Margins”. In: *Available at SSRN* 4136866. DOI: [10.2139/ssrn.4136866](https://doi.org/10.2139/ssrn.4136866).
- Bems, Rudolfs, Francesca Caselli, Francesco Grigoli, and Bertrand Gruss (2021). “Expectations’ anchoring and inflation persistence”. In: *Journal of International Economics* 132, p. 103516. ISSN: 0022-1996. DOI: [10.1016/j.jinteco.2021.103516](https://doi.org/10.1016/j.jinteco.2021.103516).

- Bernanke, Ben S. (Feb. 2004). *The Great Moderation*. Remarks at the meetings of the Eastern Economic Association, Washington, DC. URL: <https://www.federalreserve.gov/boarddocs/speeches/2004/20040220/>.
- Bernanke, Ben S. (July 2007). *Inflation expectations and inflation forecasting*. Speech at the NBER Summer Institute, Cambridge, Massachusetts. URL: <https://www.federalreserve.gov/newsevents/speech/bernanke20070710a.htm>.
- Bianchi, Francesco, Leonardo Melosi, and Matthias Rottner (2021). “Hitting the elusive inflation target”. In: *Journal of Monetary Economics* 124, pp. 107–122. ISSN: 0304-3932. DOI: [10.1016/j.jmoneco.2021.10.003](https://doi.org/10.1016/j.jmoneco.2021.10.003).
- Boero, Gianna, Jeremy Smith, and Kenneth F. Wallis (2008). “Uncertainty and Disagreement in Economic Prediction: The Bank of England Survey of External Forecasters”. In: *Economic Journal* 118.530, pp. 1107–1127. DOI: [10.1111/j.1468-0297.2008.02162.x](https://doi.org/10.1111/j.1468-0297.2008.02162.x).
- Borio, Claudio, Leonardo Gambacorta, and Boris Hofmann (2017). “The influence of monetary policy on bank profitability”. en. In: *International Finance* 20.1. _eprint: <https://onlinelibrary.wiley.com/doi/abs/10.1111/infi.12104> (visited on 04/01/2024). pp. 48–63. ISSN: 1468-2362. DOI: [10.1111/infi.12104](https://doi.org/10.1111/infi.12104). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/infi.12104> (visited on 04/01/2024).
- Breeden, Douglas T and Robert H Litzenberger (1978). “Prices of state-contingent claims implicit in option prices”. In: *Journal of business*, pp. 621–651.
- Bundick, Brent and A. Lee Smith (2025). “Did the Federal Reserve Break the Phillips Curve? Theory and Evidence of Anchoring Inflation Expectations”. In: *Review of Economics and Statistics* 107.5, pp. 1310–1326. DOI: [10.1162/rest_a_01357](https://doi.org/10.1162/rest_a_01357).
- Buseti, Fabio (Aug. 2017). “Quantile Aggregation of Density Forecasts”. en. In: *Oxford Bulletin of Economics and Statistics* 79.4, pp. 495–512. ISSN: 0305-9049, 1468-0084. DOI: [10.1111/obes.12163](https://doi.org/10.1111/obes.12163). URL: <https://onlinelibrary.wiley.com/doi/10.1111/obes.12163> (visited on 03/25/2024).

- Canals-Cerda, José J. and Brian Jonghwan Lee (2021). “COVID-19 and Auto Loan Origination Trends”. en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: [10.2139/ssrn.3897908](https://doi.org/10.2139/ssrn.3897908). URL: <https://www.ssrn.com/abstract=3897908> (visited on 02/16/2026).
- Capistrán, Carlos (2008). “Bias in Federal Reserve inflation forecasts: Is the Federal Reserve irrational or just cautious?” In: *Journal of Monetary Economics* 55.8, pp. 1415–1427. ISSN: 0304-3932. DOI: [10.1016/j.jmoneco.2008.09.011](https://doi.org/10.1016/j.jmoneco.2008.09.011).
- Capistrán, Carlos and Manuel Ramos-Francia (2010). “Does Inflation Targeting Affect the Dispersion of Inflation Expectations?” In: *Journal of Money, Credit and Banking* 42.1, pp. 113–134. DOI: [10.1111/j.1538-4616.2009.00280.x](https://doi.org/10.1111/j.1538-4616.2009.00280.x).
- Carriero, Andrea, Todd E Clark, and Massimiliano Marcellino (2020). *Nowcasting Tail Risks to Economic Activity with Many Indicators*. Tech. rep. Federal Reserve Bank of Cleveland. DOI: [10.26509/frbc-wp-202013r2](https://doi.org/10.26509/frbc-wp-202013r2).
- Carriero, Andrea, Todd E Clark, and Massimiliano Marcellino (2024). “Specification choices in quantile regression for empirical macroeconomics”. In: *Journal of Applied Econometrics*. DOI: [10.1002/jae.3099](https://doi.org/10.1002/jae.3099).
- Carvalho, Carlos, Stefano Eusepi, Emanuel Moench, and Bruce Preston (Jan. 2023). “Anchored Inflation Expectations”. In: *American Economic Journal: Macroeconomics* 15.1, pp. 1–47. DOI: [10.1257/mac.20200080](https://doi.org/10.1257/mac.20200080).
- Chang, Minsu, Xiaohong Chen, and Frank Schorfheide (2024). “Heterogeneity and Aggregate Fluctuations”. In: *Journal of Political Economy* 132.6, pp. 2045–2106. DOI: [10.1086/729444](https://doi.org/10.1086/729444).
- Chen, Qingqing, Dennis Glennon, and Amos Golan (2020). “Estimating Conditional Mortgage Delinquency Transition Matrices”. en. In: *OCC Working Paper Series*.
- Clarida, Richard, Jordi Galí, and Mark Gertler (2000). “Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory”. In: *Quarterly Journal of Economics* 115.1. Also available as NBER Working Paper 6442, pp. 147–180. DOI: [10.1162/003355300554692](https://doi.org/10.1162/003355300554692).

- Clements, Michael P. (2014). “Forecast Uncertainty—Ex Ante and Ex Post: U.S. Inflation and Output Growth”. In: *Journal of Business & Economic Statistics* 32.2, pp. 206–216. DOI: [10.1080/07350015.2013.859618](https://doi.org/10.1080/07350015.2013.859618).
- Clements, Michael P. and Shixuan Wang (2024). *Do Professionals’ Inflation Forecasts Incorporate the Beliefs of Others? A Functional Data Approach*. Working Paper 4357012. SSRN. DOI: [10.2139/ssrn.4357012](https://doi.org/10.2139/ssrn.4357012).
- Cline, Alexander and Robert W. Rich (Jan. 2026). *How Anchored Are Short-Run Inflation Expectations Today? A Look at What Consumers and Forecasters Are Telling Us*. Economic Commentary 2026-01. Federal Reserve Bank of Cleveland. DOI: [10.26509/frbc-ec-202601](https://doi.org/10.26509/frbc-ec-202601).
- Covas, Francisco, Ben Rump, and Egon Zakrajšek (July 2014). “Stress-testing US bank holding companies: A dynamic panel quantile regression approach”. en. In: *International Journal of Forecasting* 30.3, pp. 691–713. ISSN: 01692070. DOI: [10.1016/j.ijforecast.2013.11.003](https://doi.org/10.1016/j.ijforecast.2013.11.003). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0169207013001623> (visited on 09/01/2023).
- Deng, Yongheng, John M. Quigley, and Robert Order (Mar. 2000). “Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options”. en. In: *Econometrica* 68.2, pp. 275–307. ISSN: 0012-9682, 1468-0262. DOI: [10.1111/1468-0262.00110](https://doi.org/10.1111/1468-0262.00110). URL: <http://doi.wiley.com/10.1111/1468-0262.00110> (visited on 02/17/2026).
- Detting, Lisa and Lauren Lambie-Hanson (Mar. 2021). “Why is the Default Rate So Low? How Economic Conditions and Public Policies Have Shaped Mortgage and Auto Delinquencies During the COVID-19 Pandemic”. en. In: *FEDS Notes* 2021.0.2854. ISSN: 2380-7172. DOI: [10.17016/2380-7172.2854](https://doi.org/10.17016/2380-7172.2854). URL: <https://www.federalreserve.gov/econres/notes/feds-notes/why-is-the-default-rate-so-low-20210304.htm> (visited on 02/17/2026).

- Di Lucido, Katherine, Anna Kovner, and Samantha Zeller (June 2017). *Low Interest Rates and Bank Profits*. en-US. URL: <https://libertystreeteconomics.newyorkfed.org/2017/06/low-interest-rates-and-bank-profits/> (visited on 03/25/2024).
- Douglas Smith, L. and Edward C. Lawrence (Sept. 1995). “Forecasting losses on a liquidating long-term loan portfolio”. en. In: *Journal of Banking & Finance* 19.6, pp. 959–985. ISSN: 03784266. DOI: [10.1016/0378-4266\(94\)00065-B](https://doi.org/10.1016/0378-4266(94)00065-B). URL: <https://linkinghub.elsevier.com/retrieve/pii/037842669400065B> (visited on 02/16/2026).
- Drechsler, Itamar, Alexi Savov, and Philipp Schnabl (June 2021). “Banking on Deposits: Maturity Transformation without Interest Rate Risk”. en. In: *The Journal of Finance* 76.3, pp. 1091–1143. ISSN: 0022-1082, 1540-6261. DOI: [10.1111/jofi.13013](https://doi.org/10.1111/jofi.13013). URL: <https://onlinelibrary.wiley.com/doi/10.1111/jofi.13013> (visited on 09/01/2023).
- Egozcue, Juan José, José Luis Díaz-Barrero, and Vera Pawlowsky-Glahn (2006). “Hilbert space of probability density functions based on Aitchison geometry”. In: *Acta Mathematica Sinica* 22.4, pp. 1175–1182.
- English, William (2002). “Interest rate risk and bank net interest margins”. In: *BIS Quarterly Review*. Publisher: Bank for International Settlements. URL: <https://econpapers.repec.org/article/bisbisqtr/0212g.htm> (visited on 03/25/2024).
- English, William, Skander J. Van Den Heuvel, and Egon Zakrajšek (Oct. 2018). “Interest rate risk and bank equity valuations”. en. In: *Journal of Monetary Economics* 98, pp. 80–97. ISSN: 03043932. DOI: [10.1016/j.jmoneco.2018.04.010](https://doi.org/10.1016/j.jmoneco.2018.04.010). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0304393218302101> (visited on 09/01/2023).
- Ennis, Huberto M and Arantxa Jarque (2021). “Bank Lending in the Time of COVID”. In.
- Fairlie, Robert W. (Nov. 2005). “An extension of the Blinder-Oaxaca decomposition technique to logit and probit models”. en. In: *Journal of Economic and Social Measurement* 30.4, pp. 305–316. ISSN: 0747-9662, 1875-8932. DOI: [10.3233/JEM-2005-0259](https://doi.org/10.3233/JEM-2005-0259). URL: <https://journals.sagepub.com/doi/10.3233/JEM-2005-0259> (visited on 02/17/2026).

- Flannery, Mark (1981). “Market Interest Rates and Commercial Bank Profitability: An Empirical Investigation”. en. In: *The Journal of Finance* 36.5. _eprint: <https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1981.tb01078.x>, pp. 1085–1101. ISSN: 1540-6261. DOI: [10.1111/j.1540-6261.1981.tb01078.x](https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1981.tb01078.x). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1981.tb01078.x> (visited on 03/25/2024).
- Galí, Jordi (2008). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton, NJ: Princeton University Press. ISBN: 978-0-691-13316-4.
- Galvao, Antonio F. (Sept. 2011). “Quantile regression for dynamic panel data with fixed effects”. en. In: *Journal of Econometrics* 164.1, pp. 142–157. ISSN: 03044076. DOI: [10.1016/j.jeconom.2011.02.016](https://linkinghub.elsevier.com/retrieve/pii/S0304407611000443). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0304407611000443> (visited on 09/01/2023).
- Gáti, Laura (2023). “Monetary policy and anchored expectations – An endogenous gain learning model”. In: *Journal of Monetary Economics* 140, S37–S47. DOI: [10.1016/j.jmoneco.2023.06.009](https://doi.org/10.1016/j.jmoneco.2023.06.009).
- Giglio, Carla, Frances Shaw, Nicolas Syrichas, and Giuseppe Cappelletti (Feb. 2021). *Stress-Testing Net Trading Income: The Case of European Banks*. en. ECB Working Paper 2021/2525. URL: <https://papers.ssrn.com/abstract=3797145> (visited on 03/25/2024).
- Gneiting, Tilmann (2011). “Quantiles as optimal point forecasts”. In: *International Journal of Forecasting* 27.2, pp. 197–207. ISSN: 0169-2070. DOI: [10.1016/j.ijforecast.2009.12.015](https://doi.org/10.1016/j.ijforecast.2009.12.015).
- Gneiting, Tilmann and Matthias Katzfuss (2014). “Probabilistic forecasting”. In: *Annual Review of Statistics and Its Application* 1.1, pp. 125–151. DOI: [10.1146/annurev-statistics-062713-085831](https://doi.org/10.1146/annurev-statistics-062713-085831).
- Gneiting, Tilmann and Adrian E Raftery (2007). “Strictly proper scoring rules, prediction, and estimation”. In: *Journal of the American statistical Association* 102.477, pp. 359–378. DOI: [10.1198/016214506000001437](https://doi.org/10.1198/016214506000001437).

- Gneiting, Tilmann and Roopesh Ranjan (2011). “Comparing density forecasts using threshold- and quantile-weighted scoring rules”. In: *Journal of Business & Economic Statistics* 29.3, pp. 411–422. DOI: [10.1198/jbes.2010.08110](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”. In: DOI: [10.2139/ssrn.52400](https://doi.org/10.2139/ssrn.52400).
- Grishchenko, Olesya, Sarah Mouabbi, and Jean-Paul Renne (2019). “Measuring Inflation Anchoring and Uncertainty: A U.S. and Euro Area Comparison”. In: *Journal of Money, Credit and Banking* 51.5, pp. 1053–1096. DOI: [10.1111/jmcb.12622](https://doi.org/10.1111/jmcb.12622).
- Gross, Marco and Javier Población (Feb. 2019). “Implications of Model Uncertainty for Bank Stress Testing”. en. In: *Journal of Financial Services Research* 55.1, pp. 31–58. ISSN: 0920-8550, 1573-0735. DOI: [10.1007/s10693-017-0275-4](https://doi.org/10.1007/s10693-017-0275-4). URL: <http://link.springer.com/10.1007/s10693-017-0275-4> (visited on 09/01/2023).
- Guerrieri, Luca and Michelle Welch (July 2012). *Can Macro Variables Used in Stress Testing Forecast the Performance of Banks?* en. SSRN Scholarly Paper. Rochester, NY. DOI: [10.2139/ssrn.2192650](https://doi.org/10.2139/ssrn.2192650). URL: <https://papers.ssrn.com/abstract=2192650> (visited on 03/25/2024).
- Guo, Zhengfeng, Yan Zhang, and Xinlei Zhao (2022). “Risks of long-term auto loans”. In: *Journal of Credit Risk*. ISSN: 17446619, 17559723. DOI: [10.21314/JCR.2022.009](https://doi.org/10.21314/JCR.2022.009). URL: <https://www.risk.net/journal-of-credit-risk/7954991/risks-of-long-term-auto-loans> (visited on 02/17/2026).
- Gürkaynak, Refet S., Andrew Levin, and Eric Swanson (Dec. 2010). “Does Inflation Targeting Anchor Long-Run Inflation Expectations? Evidence from the U.S., UK, and Sweden”. In: *Journal of the European Economic Association* 8.6, pp. 1208–1242. DOI: [10.1111/j.1542-4774.2010.tb00553.x](https://doi.org/10.1111/j.1542-4774.2010.tb00553.x).
- Hanweck, Gerald A. and Lisa H. Ryu (Jan. 2005). *The Sensitivity of Bank Net Interest Margins and Profitability To Credit, Interest-Rate, and Term-Structure Shocks Across Bank Product Specializations*. en. FDIC Working Paper 05-02. Rochester, NY. DOI: [10.2139/ssrn.848888](https://doi.org/10.2139/ssrn.848888).

- 2139/ssrn.886727. URL: <https://papers.ssrn.com/abstract=886727> (visited on 03/25/2024).
- Haubrich, Joseph G. and Tristan Young (2019). “Trends in the Noninterest Income of Banks”. en. In: *Federal Reserve Bank of Cleveland Economic Commentary* 2019-14. Publisher: Federal Reserve Bank of Cleveland. DOI: [10.26509/frbc-ec-201914](https://ideas.repec.org/a/fip/fedcec/00106.html). URL: <https://ideas.repec.org/a/fip/fedcec/00106.html> (visited on 03/25/2024).
- Heitfield, Erik and Tarun Sabarwal (Dec. 2004). “What Drives Default and Prepayment on Subprime Auto Loans?” en. In: *The Journal of Real Estate Finance and Economics* 29.4, pp. 457–477. ISSN: 0895-5638, 1573-045X. DOI: [10.1023/B:REAL.0000044023.02636.e6](https://link.springer.com/10.1023/B:REAL.0000044023.02636.e6). URL: <https://link.springer.com/10.1023/B:REAL.0000044023.02636.e6> (visited on 02/16/2026).
- Hirtle, Beverly, Anna Kovner, James Vickery, and Meru Bhanot (Aug. 2016). “Assessing financial stability: The Capital and Loss Assessment under Stress Scenarios (CLASS) model”. en. In: *Journal of Banking & Finance* 69, S35–S55. ISSN: 03784266. DOI: [10.1016/j.jbankfin.2015.09.021](https://linkinghub.elsevier.com/retrieve/pii/S0378426615002940). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0378426615002940> (visited on 09/01/2023).
- Ho, Thomas S. Y. and Anthony Saunders (Nov. 1981). “The Determinants of Bank Interest Margins: Theory and Empirical Evidence”. en. In: *The Journal of Financial and Quantitative Analysis* 16.4, p. 581. ISSN: 00221090. DOI: [10.2307/2330377](https://www.jstor.org/stable/2330377?origin=crossref). URL: <https://www.jstor.org/stable/2330377?origin=crossref> (visited on 09/01/2023).
- Hron, Karel, Alessandra Menafoglio, Matthias Templ, Klára Hrzová, and Peter Filzmoser (2016). “Simplicial principal component analysis for density functions in Bayes spaces”. In: *Computational Statistics & Data Analysis* 94, pp. 330–350.
- Huber, Florian, Massimiliano Marcellino, and Tommaso Tornese (2024a). “The Distributional Effects of Economic Uncertainty”. In: *arXiv preprint arXiv:2411.12655*.
- Huber, Florian, Massimiliano Marcellino, and Tommaso Tornese (2024b). “The Distributional Effects of Economic Uncertainty”. In: *arXiv preprint arXiv:2411.12655*.

- Jochmann, Markus, Gary Koop, and Simon M. Potter (2010). “Modeling the dynamics of inflation compensation”. In: *Journal of Empirical Finance* 17.1, pp. 157–167. ISSN: 0927-5398. DOI: [10.1016/j.jempfin.2009.08.002](https://doi.org/10.1016/j.jempfin.2009.08.002).
- Katcher, Bradley, Geng Li, Alvaro Mezza, and Steve Ramos (July 2024). “One Month Longer, One Month Later? Prepayments in the Auto Loan Market”. en. In: *Finance and Economics Discussion Series* 2024-056, pp. 1–66. ISSN: 1936-2854, 2767-3898. DOI: [10.17016/feds.2024.056](https://doi.org/10.17016/feds.2024.056). URL: <https://www.federalreserve.gov/econres/feds/one-month-longer-one-month-later-prepayments-in-the-auto-loan-market.htm> (visited on 02/16/2026).
- Kitsul, Yuriy and Jonathan H Wright (2013). “The economics of options-implied inflation probability density functions”. In: *Journal of Financial Economics* 110.3, pp. 696–711.
- Koenker, Roger (Oct. 2004). “Quantile regression for longitudinal data”. en. In: *Journal of Multivariate Analysis* 91.1, pp. 74–89. ISSN: 0047259X. DOI: [10.1016/j.jmva.2004.05.006](https://doi.org/10.1016/j.jmva.2004.05.006). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0047259X04001113> (visited on 09/01/2023).
- Kozumi, Hideo and Genya Kobayashi (Nov. 2011). “Gibbs sampling methods for Bayesian quantile regression”. In: *Journal of Statistical Computation and Simulation* 81.11. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/00949655.2010.496117>, pp. 1565–1578. ISSN: 0094-9655. DOI: [10.1080/00949655.2010.496117](https://doi.org/10.1080/00949655.2010.496117). URL: <https://doi.org/10.1080/00949655.2010.496117> (visited on 03/25/2024).
- Kumar, Saten, Hassan Afrouzi, Olivier Coibion, and Yuriy Gorodnichenko (2015). “Inflation Targeting Does Not Anchor Inflation Expectations: Evidence from Firms in New Zealand”. In: *Brookings Papers on Economic Activity* 46.2 (Fall), pp. 151–225. URL: <https://ideas.repec.org/a/bin/bpeajo/v46y2015i2015-02p151-225.html>.
- 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”. en. In: *Federal Reserve Bank of Kansas City, Economic Review*. DOI: [10.18651/ER/v109n1LaliberteSengupta](https://doi.org/10.18651/ER/v109n1LaliberteSengupta).
- Lancaster, Tony (1990). “The econometric analysis of transition data”. In.
- Lucas, Robert E. (1976). “Econometric policy evaluation: A critique”. en. In: *Carnegie-Rochester Conference Series on Public Policy* 1, pp. 19–46. ISSN: 01672231. DOI: [10.1016/S0167-2231\(76\)80003-6](https://doi.org/10.1016/S0167-2231(76)80003-6). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0167223176800036> (visited on 02/17/2026).
- Meeks, Raphael and Francesca Monti (2023). “Heterogeneous Beliefs and the Phillips Curve”. In: *Journal of Monetary Economics* 140. Carnegie-Rochester-NYU Conference Series on Public Policy, S86–S104. DOI: [10.1016/j.jmoneco.2023.10.012](https://doi.org/10.1016/j.jmoneco.2023.10.012).
- Mincer, Jacob A and Victor Zarnowitz (1969). “The evaluation of economic forecasts”. In: *Economic forecasts and expectations: Analysis of forecasting behavior and performance*. NBER, pp. 3–46. URL: <https://www.nber.org/system/files/chapters/c1214/c1214.pdf>.
- Müller, Ulrich K (2013). “Risk of Bayesian inference in misspecified models, and the sandwich covariance matrix”. In: *Econometrica* 81.5, pp. 1805–1849.
- Naggert, Kristoph, Robert W. Rich, and Joseph Tracy (July 2023). *The Anchoring of US Inflation Expectations Since 2012*. Economic Commentary 2023-11. Federal Reserve Bank of Cleveland. DOI: [10.26509/frbc-ec-202311](https://doi.org/10.26509/frbc-ec-202311).
- Nautz, Dieter and Till Strohsal (2015). “Are US inflation expectations re-anchored?” In: *Economics Letters* 127, pp. 6–9. ISSN: 0165-1765. DOI: [10.1016/j.econlet.2014.12.023](https://doi.org/10.1016/j.econlet.2014.12.023).
- Newey, Whitney K. and Kenneth D. West (1987). “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”. In: *Econometrica* 55.3, pp. 703–708. ISSN: 00129682, 14680262. URL: <http://www.jstor.org/stable/1913610> (visited on 02/12/2026).

- Nickell, Stephen (1981). “Biases in Dynamic Models with Fixed Effects”. In: *Econometrica* 49.6. Publisher: [Wiley, Econometric Society], pp. 1417–1426. ISSN: 0012-9682. DOI: [10.2307/1911408](https://doi.org/10.2307/1911408). URL: <https://www.jstor.org/stable/1911408> (visited on 03/25/2024).
- Park, Trevor and George Casella (June 2008). “The Bayesian Lasso”. In: *Journal of the American Statistical Association* 103.482. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1198/016214508000000337>. pp. 681–686. ISSN: 0162-1459. DOI: [10.1198/016214508000000337](https://doi.org/10.1198/016214508000000337). URL: <https://doi.org/10.1198/016214508000000337> (visited on 03/25/2024).
- Petersen, Alexander and Hans-Georg Müller (2016). “Functional data analysis for density functions by transformation to a Hilbert space”. In:
- Reis, Ricardo (2021). “Losing the Inflation Anchor”. In: *Brookings Papers on Economic Activity* Fall, pp. 307–379. DOI: [10.1353/eca.2022.0004](https://doi.org/10.1353/eca.2022.0004).
- Samuelson, Paul A (1945). “The effect of interest rate increases on the banking system”. In: *American economic review* 35.1, pp. 16–27. URL: <https://www.jstor.org/stable/1810106>.
- Sengupta, Rajdeep and Fei Xue (2022). “Do Net Interest Margins for Small and Large Banks Vary Differently with Interest Rates?” en. In: *Federal Reserve Bank of Kansas City, Economic Review*. DOI: [10.18651/ER/v107n1SenguptaXue](https://doi.org/10.18651/ER/v107n1SenguptaXue).
- Smith, L. Douglas and Baiqiang Jin (Oct. 2007). “Modeling exposure to losses on automobile leases”. en. In: *Review of Quantitative Finance and Accounting* 29.3, pp. 241–266. ISSN: 0924-865X, 1573-7179. DOI: [10.1007/s11156-007-0032-0](https://doi.org/10.1007/s11156-007-0032-0). URL: <http://link.springer.com/10.1007/s11156-007-0032-0> (visited on 02/16/2026).
- Smith, L. Douglas, Susan M. Sanchez, and Edward C. Lawrence (June 1996). “A Comprehensive Model for Managing Credit Risk on Home Mortgage Portfolios”. en. In: *Decision Sciences* 27.2, pp. 291–317. ISSN: 0011-7315, 1540-5915. DOI: [10.1111/j.1540-5915.1996.tb01719.x](https://doi.org/10.1111/j.1540-5915.1996.tb01719.x). URL: <http://doi.wiley.com/10.1111/j.1540-5915.1996.tb01719.x> (visited on 02/16/2026).

- Talská, Renáta, Alessandra Menafooglio, Jitka Machalová, Karel Hron, and E Fišerová (2018). “Compositional regression with functional response”. In: *Computational Statistics & Data Analysis* 123, pp. 66–85.
- Tibshirani, Robert (Jan. 1996). “Regression Shrinkage and Selection Via the Lasso”. en. In: *Journal of the Royal Statistical Society: Series B Methodology* 58.1, pp. 267–288. ISSN: 00359246. DOI: [10 . 1111 / j . 2517 - 6161 . 1996 . tb02080 . x](https://doi.org/10.1111/j.2517-6161.1996.tb02080.x). URL: [https : / / onlinelibrary . wiley . com / doi / 10 . 1111 / j . 2517 - 6161 . 1996 . tb02080 . x](https://onlinelibrary.wiley.com/doi/10.1111/j.2517-6161.1996.tb02080.x) (visited on 09/01/2023).
- Van den Boogaart, Karl Gerald, Juan José Egozcue, and Vera Pawlowsky-Glahn (2014). “Bayes hilbert spaces”. In: *Australian & New Zealand Journal of Statistics* 56.2, pp. 171–194.
- Williams, Emily (2020). “Heterogeneity in Net-Interest Income Exposure to Interest Rate Risk and Non-Interest Expense Adjustment”. en. In.
- Woodford, Michael (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.
- Wu, Deming, Ming Fang, and Qing Wang (Dec. 2018). “An empirical study of bank stress testing for auto loans”. en. In: *Journal of Financial Stability* 39, pp. 79–89. ISSN: 15723089. DOI: [10 . 1016 / j . jfs . 2018 . 09 . 005](https://doi.org/10.1016/j.jfs.2018.09.005). URL: [https : / / linkinghub . elsevier . com / retrieve / pii / S1572308918303139](https://linkinghub.elsevier.com/retrieve/pii/S1572308918303139) (visited on 02/16/2026).
- Yu, Keming and Rana A. Moyeed (Oct. 2001). “Bayesian quantile regression”. In: *Statistics & Probability Letters* 54.4, pp. 437–447. ISSN: 0167-7152. DOI: [10 . 1016 / S0167 - 7152 \(01\) 00124 - 9](https://doi.org/10.1016/S0167-7152(01)00124-9). URL: [https : / / www . sciencedirect . com / science / article / pii / S0167715201001249](https://www.sciencedirect.com/science/article/pii/S0167715201001249) (visited on 03/25/2024).
- Zarruk, Emilio R and Jeff Madura (1992). “Optimal bank interest margin under capital regulation and deposit insurance”. In: *Journal of Financial and Quantitative analysis* 27.1, pp. 143–149. DOI: [10 . 2307 / 2331303](https://doi.org/10.2307/2331303).

Zou, Hui (Dec. 2006). “The Adaptive Lasso and Its Oracle Properties”. In: *Journal of the American Statistical Association* 101.476. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1198/0162145060000000735>. URL: <https://doi.org/10.1198/0162145060000000735> (visited on 03/25/2024).

4.10 Tables

Table 4.1: Analysis variables used in multinomial logit.

Variable	Description	Source
Loan Age	Loan age in months.	FRBNY Consumer Credit Panel / Equifax Data
Loan Age (squared)	Loan age in months (squared).	FRBNY Consumer Credit Panel / Equifax Data
Risk Score	Equifax risk score group at origination. One of <560, 560-619, 620-659, 660-719, 720-760, 760+	FRBNY Consumer Credit Panel / Equifax Data
Lender	Lender type. One of auto finance, bank, or credit union.	FRBNY Consumer Credit Panel / Equifax Data
Borrower Age	Borrower age (years).	FRBNY Consumer Credit Panel / Equifax Data
Joint Account	Equal to 1 if tradeline is a joint account, 0 otherwise.	FRBNY Consumer Credit Panel / Equifax Data
First Mortgage Flag (t-1)	Equal to 1 if borrower is a first mortgage holder as of the prior period, 0 otherwise.	FRBNY Consumer Credit Panel / Equifax Data
Ln(Balance) (t-1)	Natural log of outstanding balance in the prior period.	FRBNY Consumer Credit Panel / Equifax Data
Annual Unemployment Change	Annual change in US state unemployment.	BLS
Annual HPI Change	Annual change in US state House Price Index	FHFA
Seasonal Flag	Equal to 1 if period is in December-May, 0 otherwise.	FRBNY Consumer Credit Panel / Equifax Data

Table 4.2: Summary statistics of analysis variables, biannual observations from 2000–2025.

Variable	N	Mean	SD	Min	P50	Max
Current	1668736.00	0.82	0.38	0.00	1.00	1.00
30-89 DPD	1668736.00	0.03	0.16	0.00	0.00	1.00
90+ DPD	1668736.00	0.01	0.11	0.00	0.00	1.00
Non-Default Exit	1668736.00	0.14	0.34	0.00	0.00	1.00
Loan Age (months)	1668736.00	24.51	17.49	0.00	21.00	108.00
Equifax Risk Score	1668736.00	690.69	104.37	280.00	710.00	841.00
Borrower Age	1668736.00	45.97	14.34	17.00	45.00	99.00
Joint Account	1668736.00	0.43	0.50	0.00	0.00	1.00
Remaining Balance	1668736.00	11005.85	7234.20	0.52	9704.76	129598.15
First Mortgage Flag	1668736.00	0.31	0.46	0.00	0.00	1.00
Annual Change in State Unemployment (p.p.)	1668736.00	-0.04	2.26	-17.47	-0.27	20.63
Annual State House Price Index Change (p.p.)	1668736.00	4.65	6.65	-30.16	4.68	30.54

Source: FRBNY Consumer Credit Panel / Equifax Data, FHFA, BLS, Authors' calculations.

Table 4.3: Current State Multinomial Logit Results

	y = Transitions from Current State		
	Current-Delinquent	Current-Exit	Current-Default
Intercept	0.000*** [0.000, 0.000]	306.610*** [306.531, 306.689]	0.032*** [0.032, 0.032]
Lender: Auto Finance	1.705*** [1.694, 1.716]	0.967*** [0.954, 0.980]	1.551*** [1.550, 1.553]
Lender: Banks	1.037*** [1.033, 1.041]	1.016** [1.001, 1.032]	0.886*** [0.885, 0.886]
Loan Age (months)	1.041*** [1.038, 1.044]	1.004*** [1.003, 1.005]	1.037*** [1.033, 1.041]
Loan Age ²	1.000*** [1.000, 1.000]	1.000*** [1.000, 1.000]	1.000*** [1.000, 1.000]
Equifax Risk Score: 720-759	3.003*** [3.002, 3.004]	0.951*** [0.938, 0.964]	1.553*** [1.553, 1.553]
Equifax Risk Score: 660-719	7.572*** [7.559, 7.585]	0.940*** [0.929, 0.951]	2.735*** [2.735, 2.736]
Equifax Risk Score: 620-659	17.505*** [17.464, 17.546]	0.953*** [0.940, 0.966]	4.905*** [4.904, 4.907]
Equifax Risk Score: <560	69.243*** [68.077, 70.429]	0.827*** [0.815, 0.839]	22.227*** [22.188, 22.265]
Borrower Age	0.995*** [0.994, 0.996]	0.997*** [0.996, 0.997]	0.993*** [0.992, 0.995]
First Mortgage Flag (t-1)	0.845*** [0.843, 0.846]	1.070*** [1.058, 1.083]	0.687*** [0.687, 0.687]
Ln(Balance) (t-1)	1.277*** [1.269, 1.284]	0.438*** [0.437, 0.439]	0.721*** [0.715, 0.727]
Joint Account	0.917*** [0.916, 0.917]	0.961*** [0.950, 0.972]	0.868*** [0.868, 0.868]
Annual Unemployment Change	1.008 [0.995, 1.021]	0.998 [0.993, 1.003]	1.089*** [1.070, 1.109]
Annual HPI Change	0.990*** [0.988, 0.992]	1.012*** [1.011, 1.013]	1.003** [1.000, 1.007]
Seasonal Effect (Jan-June = 1)	0.821*** [0.821, 0.822]	1.011** [1.001, 1.022]	0.788*** [0.788, 0.789]
Num.Obs.	1181077		
R2	0.110		
R2 Adj.	0.110		
Pseudo-R2	0.11		
AUC	0.71		

Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations. Model estimated on sample data from 2000-2019. Coefficients represent relative risks. *, **, *** represent statistical significance at 0.1, 0.05, and 0.01 levels, respectively. 95 percent confidence intervals shown in brackets. Standard errors are clustered at the tradeline level.

Table 4.4: Delinquent State Multinomial Logit Results

	y = Transitions from Delinquent State		
	Delinquent-Delinquent	Delinquent-Exit	Delinquent-Default
Intercept	0.263*** [0.262, 0.263]	379.628*** [378.849, 380.410]	133.403*** [133.181, 133.627]
Lender: Auto Finance	1.086*** [1.045, 1.129]	0.818*** [0.789, 0.847]	1.096*** [1.050, 1.144]
Lender: Banks	0.866*** [0.842, 0.891]	1.113*** [1.083, 1.144]	0.929*** [0.901, 0.957]
Loan Age (months)	0.996 [0.989, 1.003]	0.987*** [0.977, 0.996]	0.942*** [0.935, 0.949]
Loan Age ²	1.000*** [1.000, 1.000]	1.000*** [1.000, 1.001]	1.001*** [1.001, 1.001]
Equifax Risk Score: 720-759	1.565*** [1.560, 1.569]	0.757*** [0.755, 0.758]	1.170*** [1.168, 1.172]
Equifax Risk Score: 660-719	1.559*** [1.467, 1.658]	0.889*** [0.873, 0.905]	1.298*** [1.269, 1.329]
Equifax Risk Score: 620-659	2.080*** [1.977, 2.190]	0.912*** [0.882, 0.943]	1.634*** [1.561, 1.711]
Equifax Risk Score: <560	2.843*** [2.722, 2.969]	0.862*** [0.816, 0.912]	2.542*** [2.428, 2.662]
Borrower Age	1.002** [1.000, 1.005]	1.000 [0.997, 1.004]	0.998 [0.996, 1.001]
First Mortgage Flag (t-1)	0.875*** [0.822, 0.930]	0.991 [0.975, 1.008]	0.541*** [0.528, 0.555]
Ln(Balance) (t-1)	1.013 [0.997, 1.030]	0.453*** [0.442, 0.464]	0.601*** [0.590, 0.611]
Joint Account	0.958* [0.914, 1.004]	0.925*** [0.896, 0.956]	0.925*** [0.878, 0.975]
Annual Unemployment Change	1.008 [0.982, 1.034]	1.057*** [1.017, 1.099]	0.994 [0.964, 1.023]
Annual HPI Change	1.001 [0.996, 1.006]	1.021*** [1.013, 1.029]	0.989*** [0.984, 0.995]
Seasonal Effect (Jan-June = 1)	0.860*** [0.821, 0.900]	0.882*** [0.854, 0.910]	0.753*** [0.715, 0.793]
Num.Obs.	29835		
R2	0.048		
R2 Adj.	0.048		
Pseudo-R2	0.05		
AUC	0.65		

Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations. Model estimated on sample data from 2000-2019. Coefficients represent relative risks. *, **, *** represent statistical significance at 0.1, 0.05, and 0.01 levels, respectively. 95 percent confidence intervals shown in brackets. Standard errors are clustered at the tradeline level.

Table 4.5: Comparison of pseudo out-of-sample 2-year cumulative forecast errors across 2007-2025.

Model	MAE	MAPE	RMSE
Binomial Logit	0.79	18	1
Markov Transition	0.44	9.9	0.5

4.11 Figures

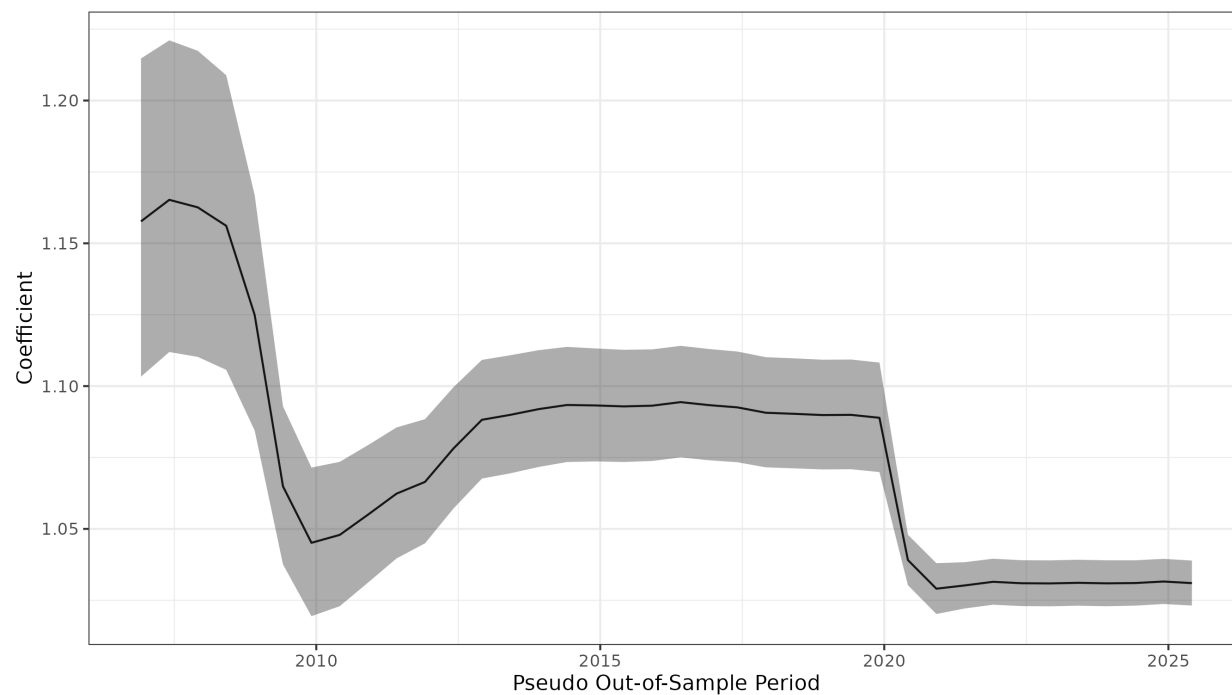


Figure 4.1: Multinomial logit estimates of US state unemployment coefficients for auto loan transitions from current to default in pseudo out-of-sample estimations, 2007-2025. Estimation variables include loan age, lender type, loan characteristics, and year-over-year changes in US state unemployment and HPI. Shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

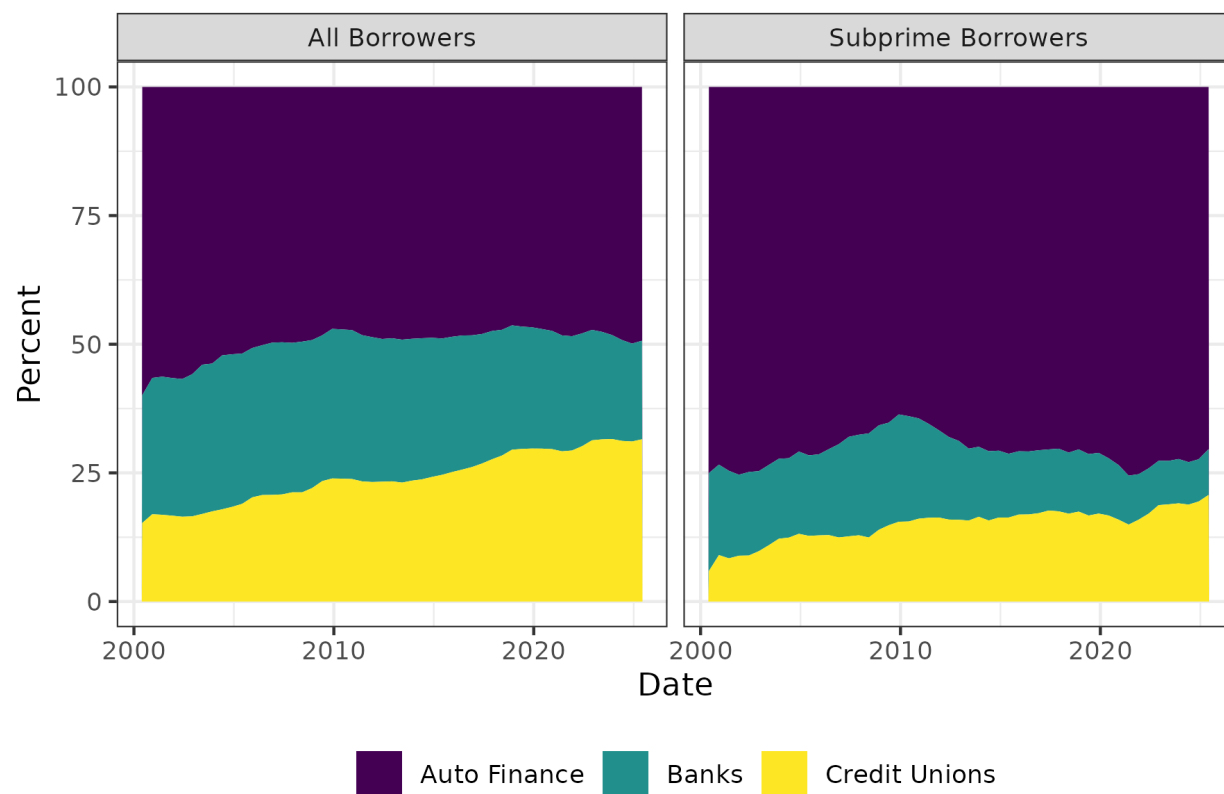


Figure 4.2: Lender types as a percent of total outstanding loans. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

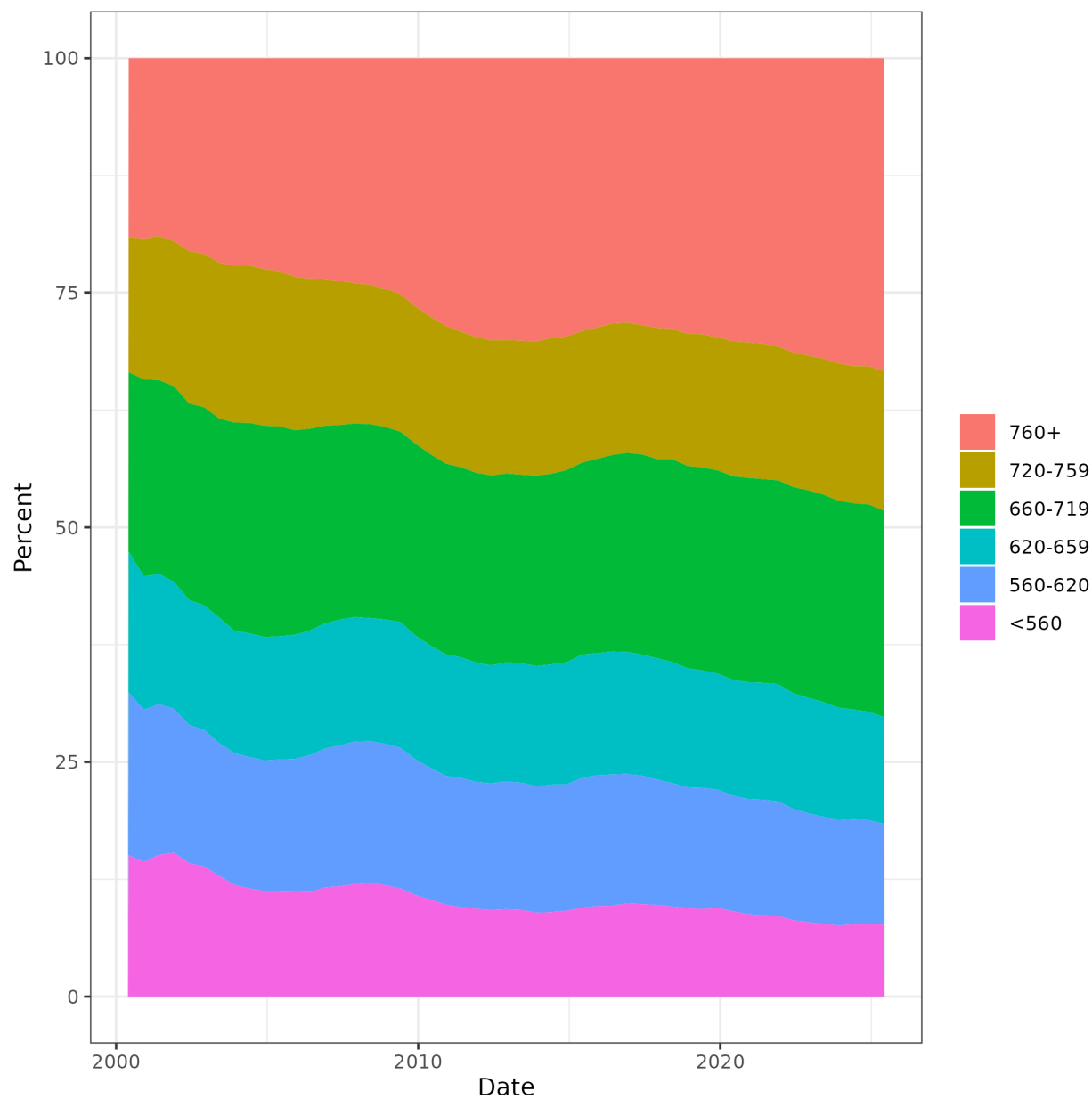


Figure 4.3: Origination Equifax Risk Scores as a percent of total outstanding loans. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

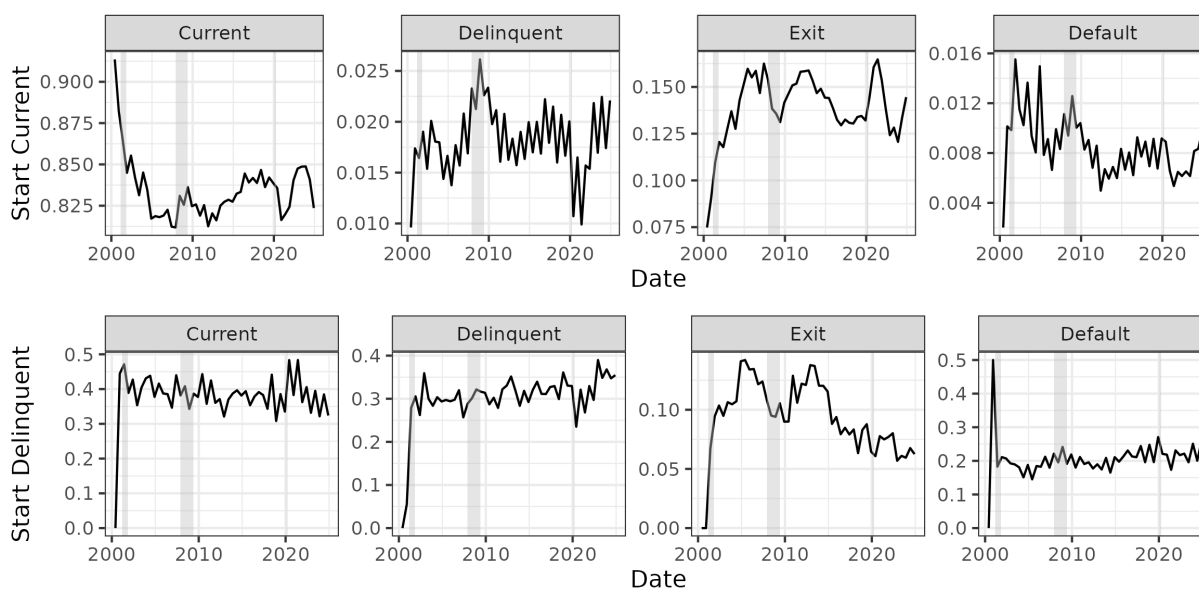


Figure 4.4: Actual transition rates by starting state, December 2000-June 2025. Note: transition rates for first few periods are distorted due to censoring at the beginning of the dataset. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

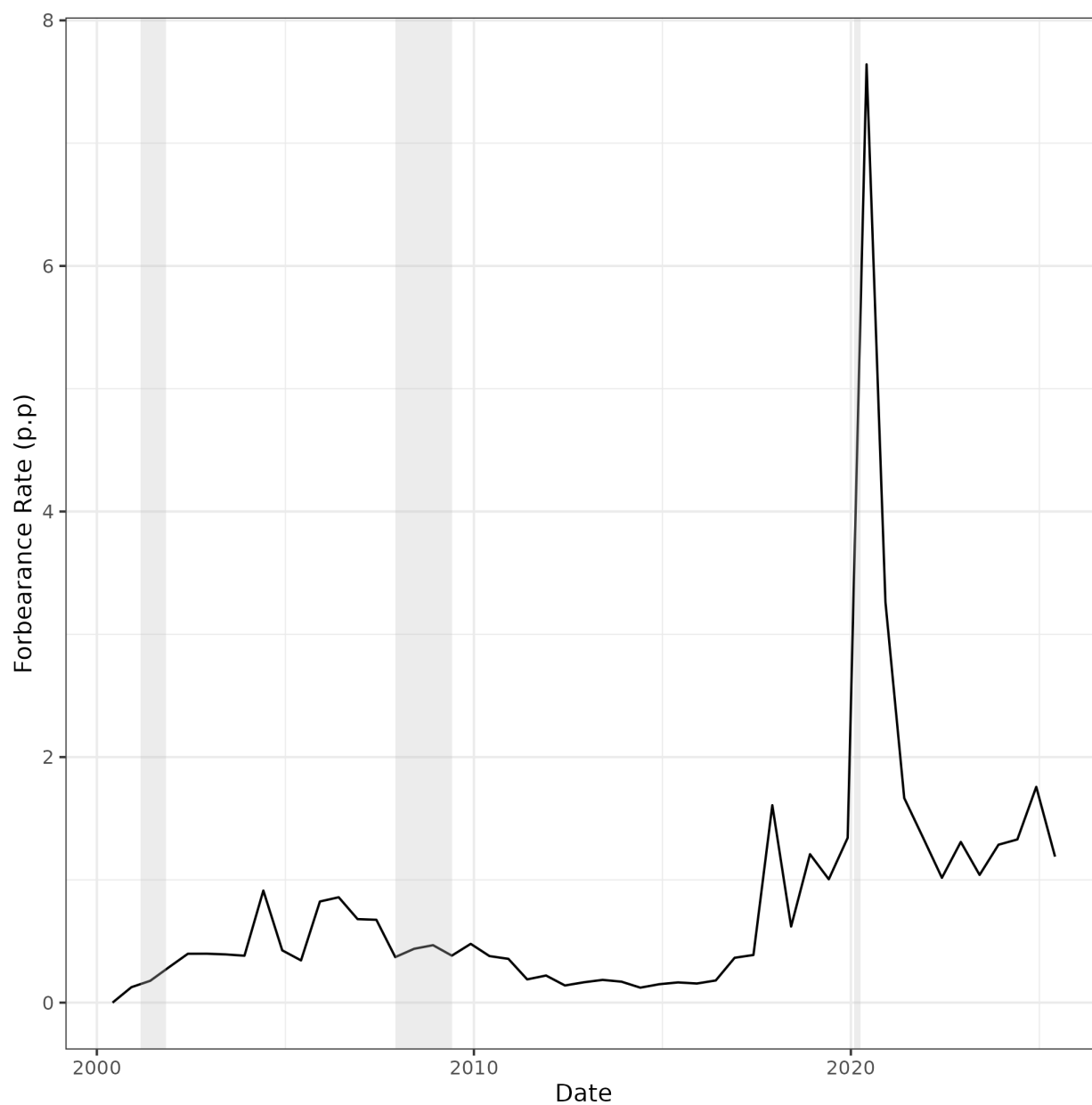


Figure 4.5: Auto loan forbearance usage, semiannual data from 2000-June 2025. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

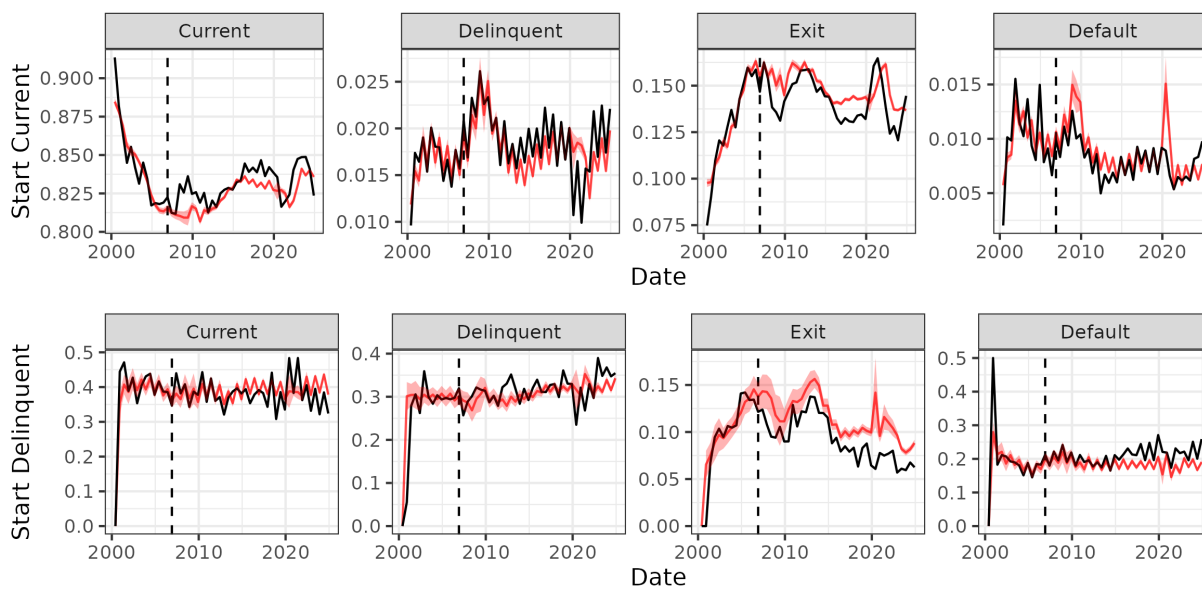


Figure 4.6: Time path of current and delinquent-state transition probabilities from pseudo out-of-sample estimation. Estimation variables include loan age, lender type, loan characteristics, and year-over-year changes in US state unemployment and HPI. The red line represents pseudo out-of-sample forecasts and the black line represents actual outcomes. Red shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Dashed line indicates December 2006, when the rolling regressions start. Pre-December 2006 estimates are in sample. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

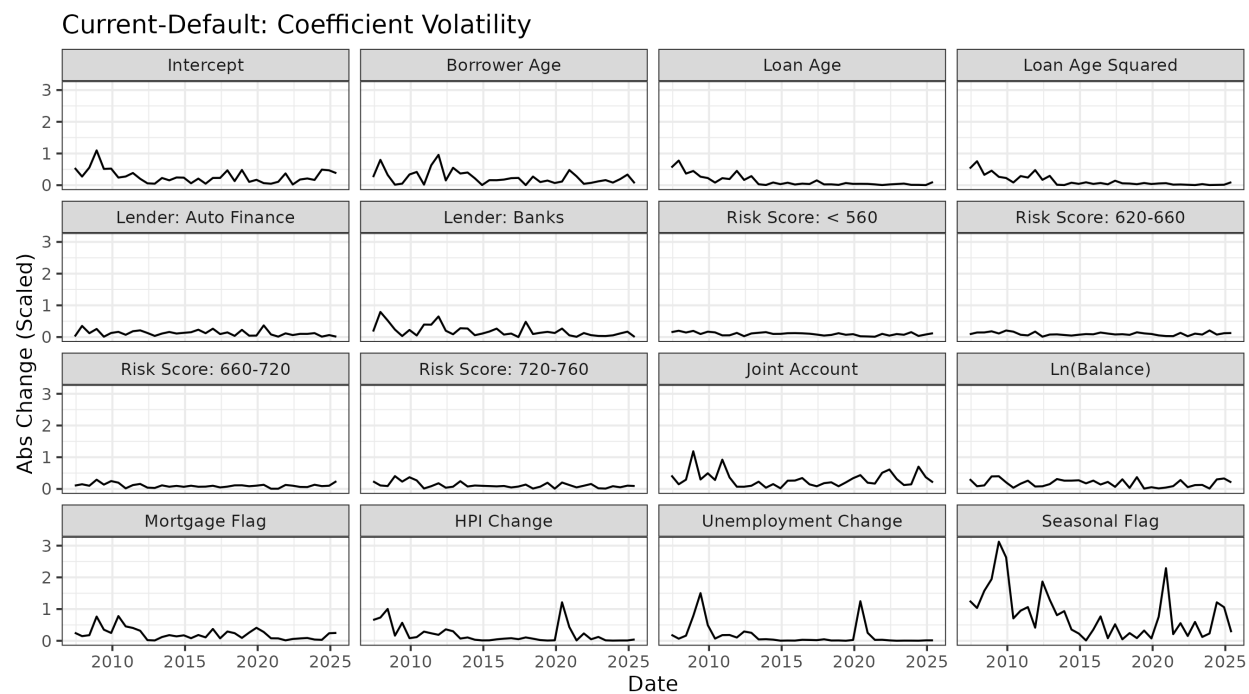


Figure 4.7: Rolling standard deviations of semiannual change in estimated relative risk coefficients for transitions from current to default. Note: Risk Score refers to the Equifax Risk Score. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

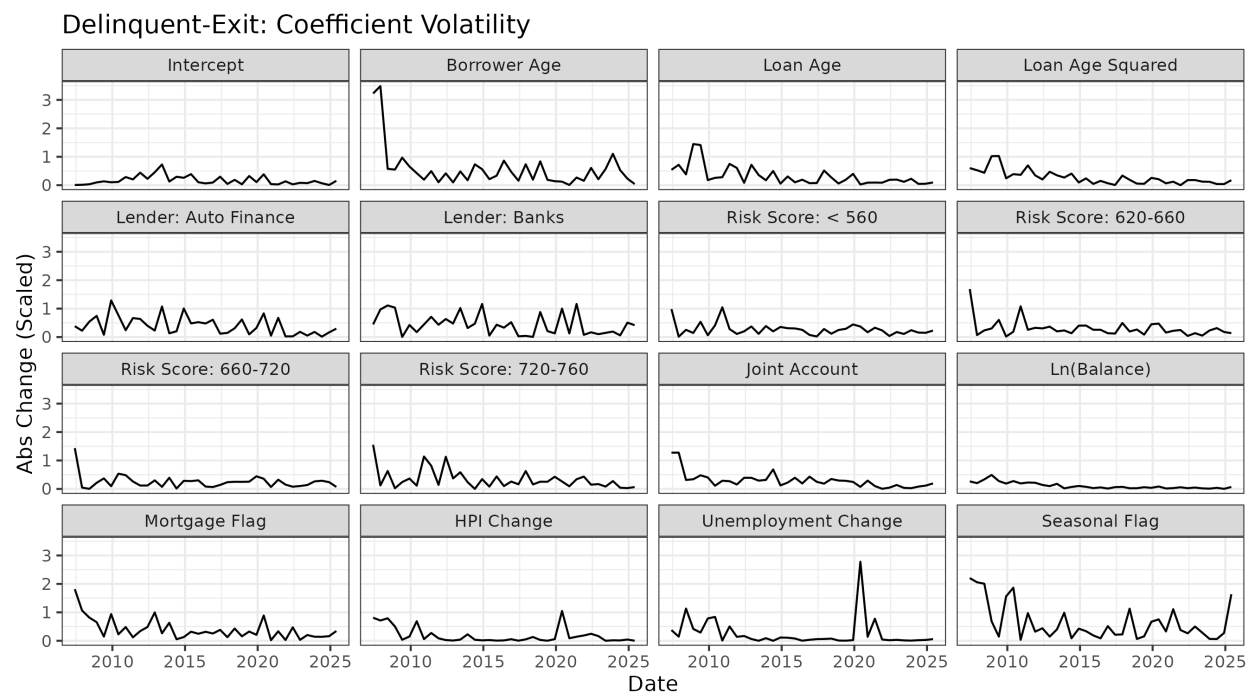


Figure 4.8: Rolling standard deviations of semiannual change in estimated relative risk coefficients for transitions from delinquent to exit. Note: Risk Score refers to the Equifax Risk Score. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

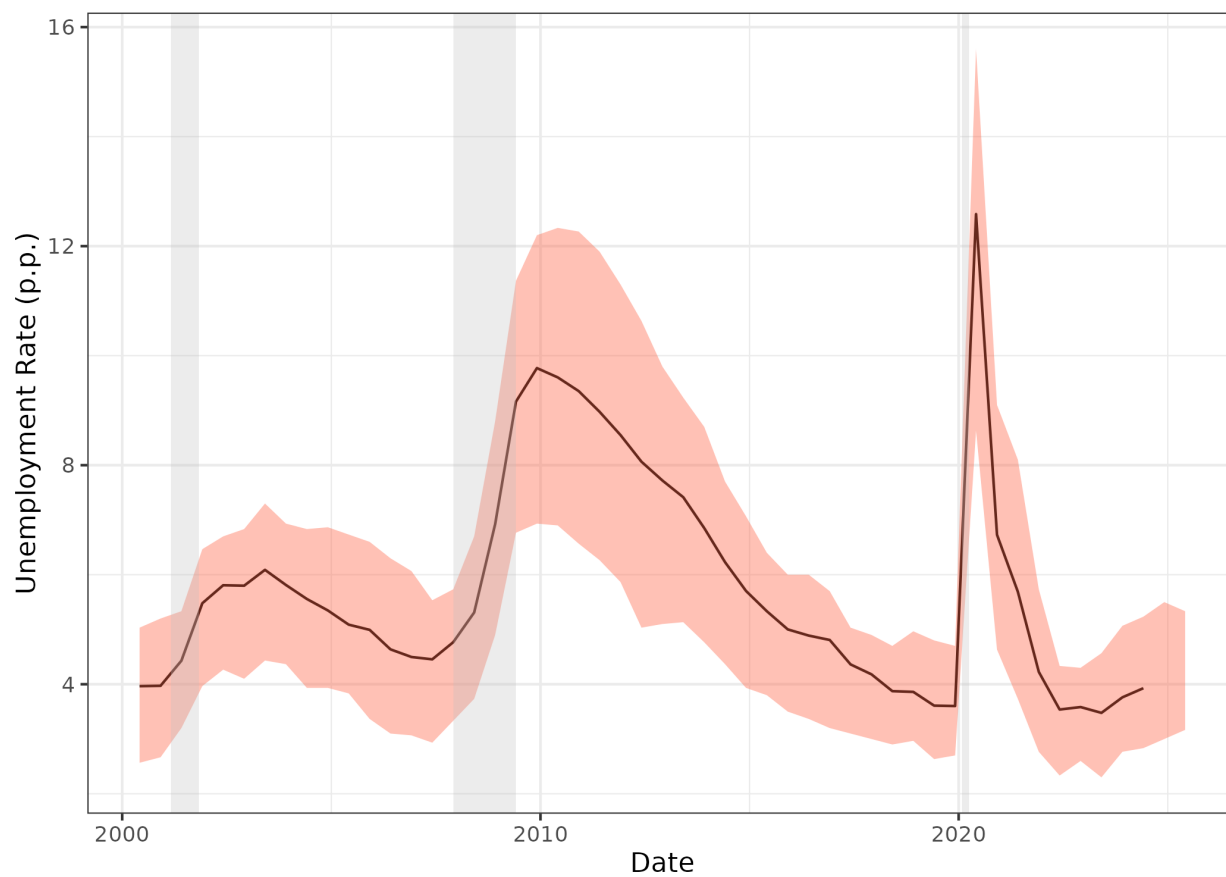


Figure 4.9: US state unemployment; semiannual data from 2000-June 2025. The solid line represents the average; the shaded area represents the 5th-95th percentiles. Source: BLS



Figure 4.10: US state house price index; semiannual data from 2000-June 2025. The solid line represents the average; the shaded area represents the 5th-95th percentiles. Source: FHFA

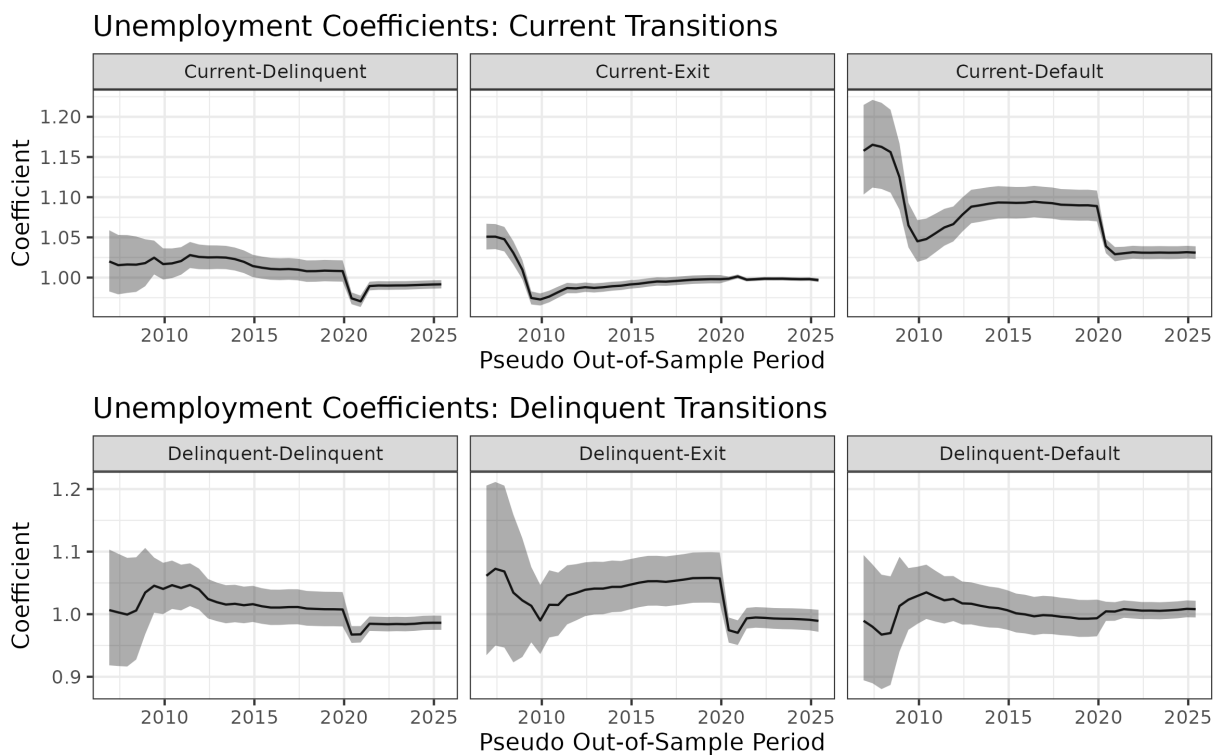


Figure 4.11: Multinomial logit estimates of US state unemployment coefficients during pseudo out-of-sample estimations, 2007-2025. Estimation variables include loan age, lender type, loan characteristics, and year-over-year changes in US state unemployment and HPI. Shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

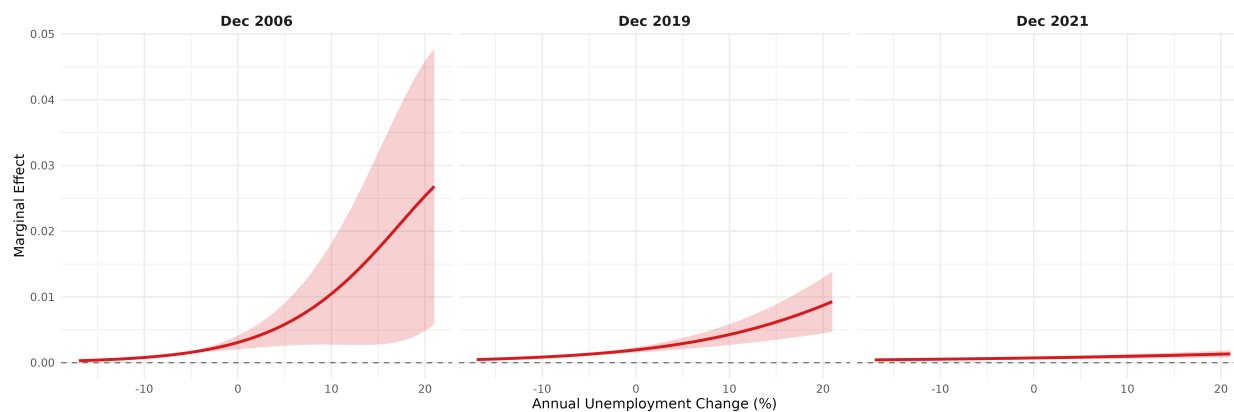


Figure 4.12: For a representative subprime borrower, the increase in the probability of transition from current to default for given year-over-year changes US state unemployment during December 2006, December 2019, and December 2021. Shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

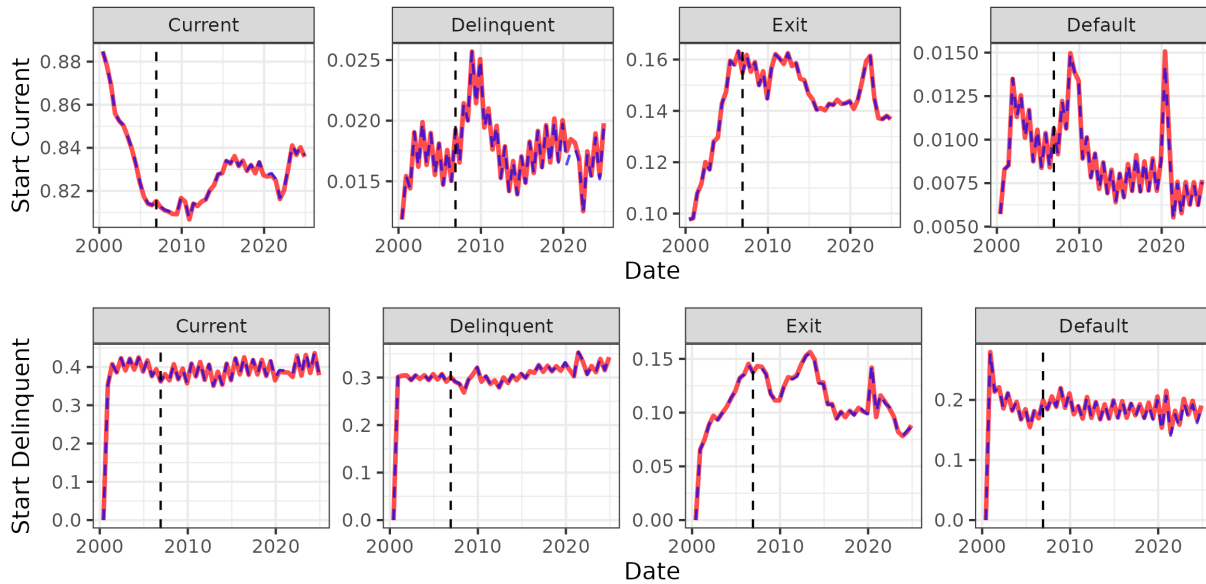


Figure 4.13: Estimated rolling regression transition probabilities on the full sample compared to the same estimation on a sample that excludes loans that enter into forbearance. The red line represents pseudo out-of-sample predictions from the full sample and the blue dashed line represents pseudo out-of-sample predictions from a sample excluding loans in forbearance. Estimation variables include loan age, lender type, loan characteristics, and year-over-year changes in US state unemployment and HPI. Dashed line indicates December 2006, when rolling regressions start. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

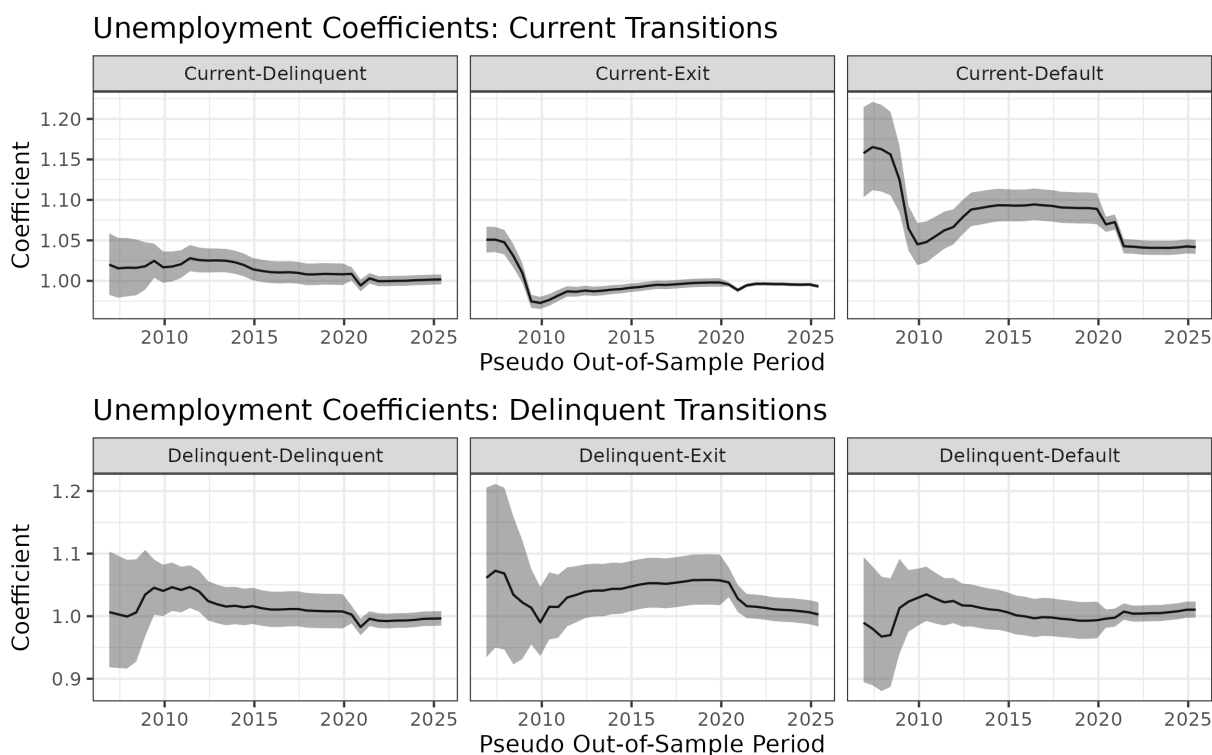


Figure 4.14: Covid Dummy Variable Results: Multinomial logit estimates of US state unemployment coefficients during pseudo out-of-sample estimations, 2012-2025. Estimation variables include loan age, lender type, loan characteristics, year-over-year changes in US state unemployment and HPI, and a dummy variable for 2000 and 2001. Shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

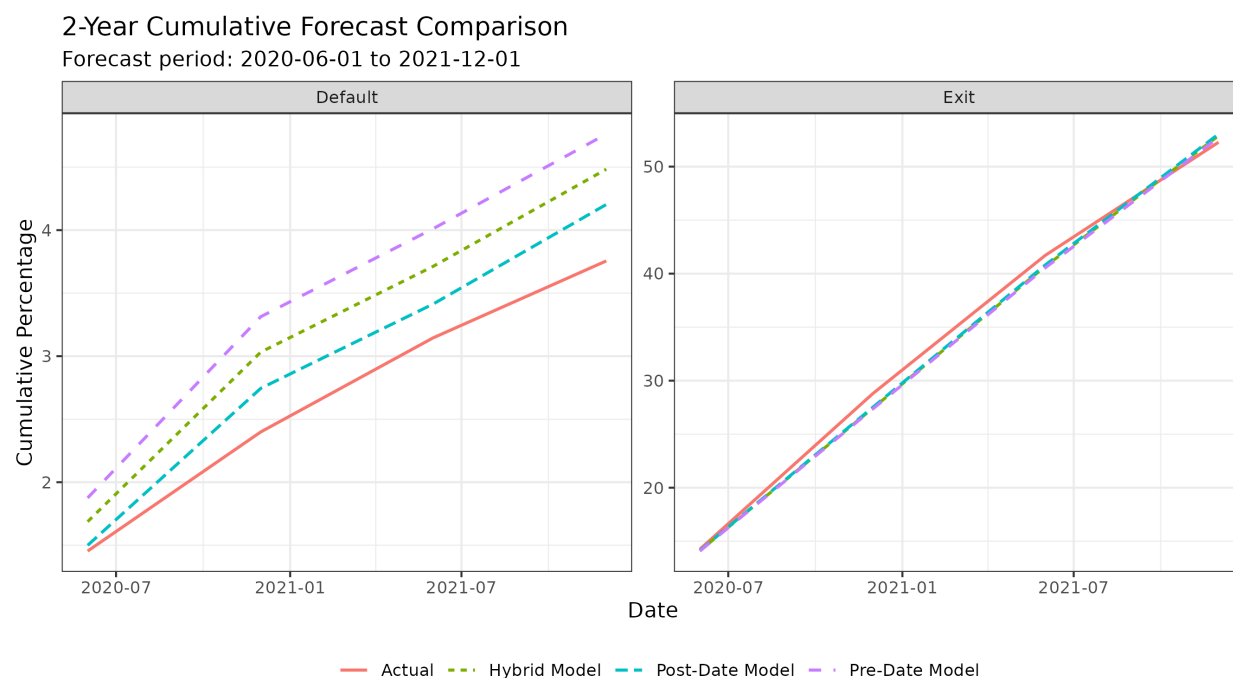


Figure 4.15: Cumulative 2-year forecasts from 2020-2021. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

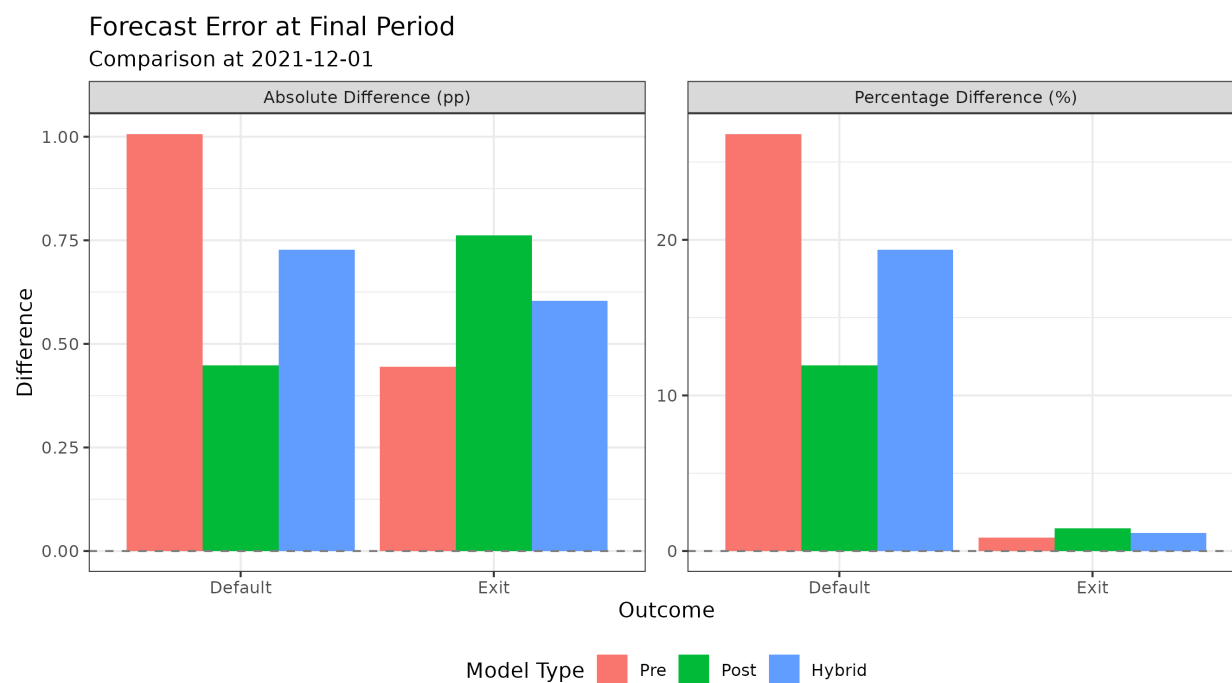


Figure 4.16: Cumulative 2-year forecast differences relative to actual outcomes. Cumulative forecasts over the period 2020-2021, differences computed as of December 2021. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

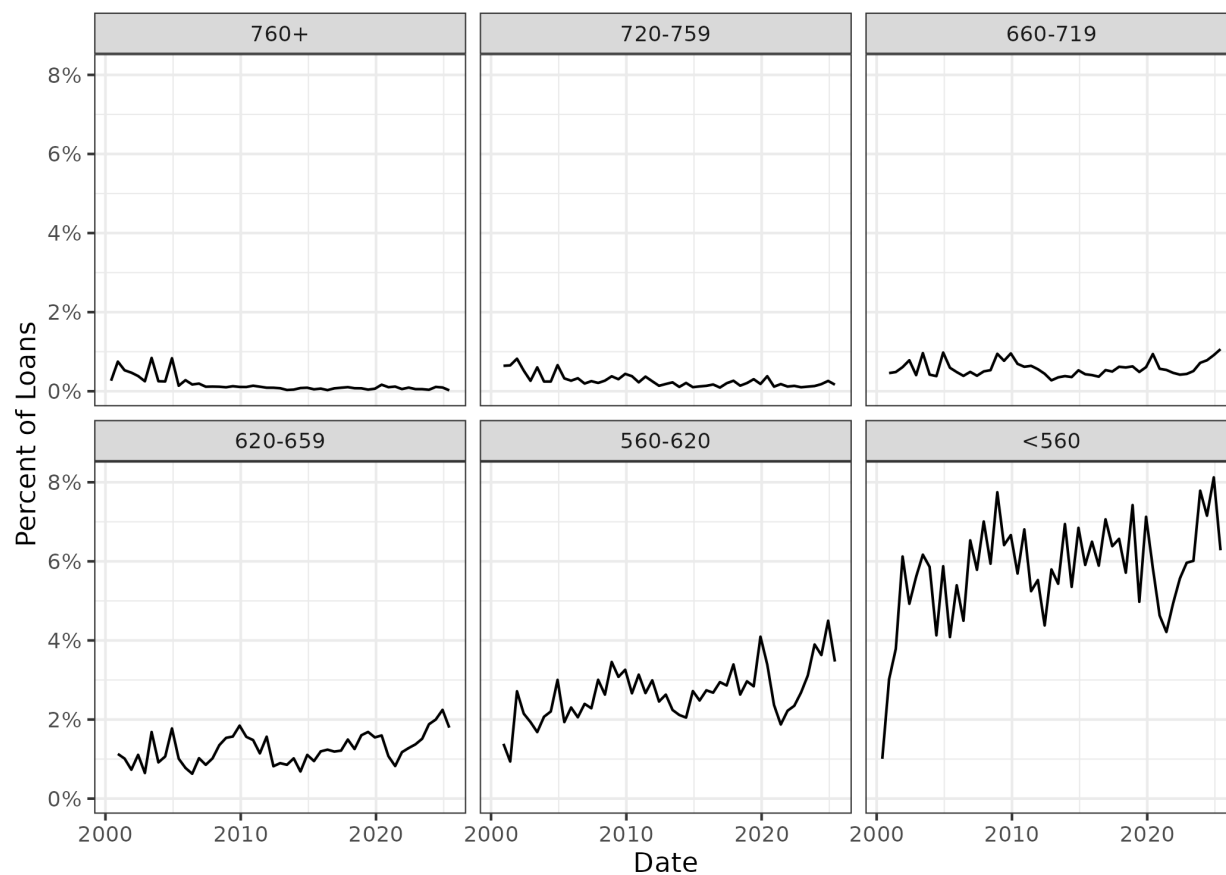


Figure 4.17: Default rates by original risk score group, December 2016-2025. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

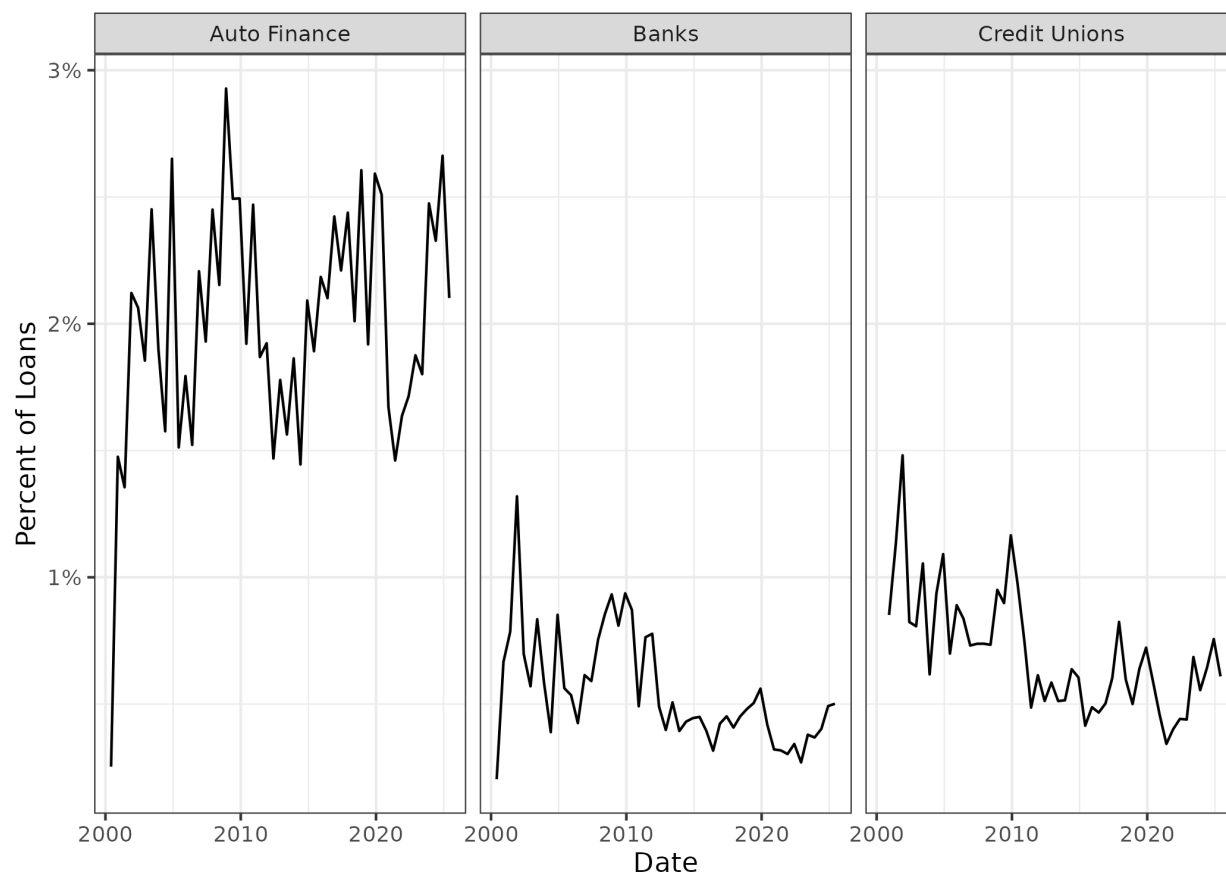


Figure 4.18: Default rates by lender type, December 2016-2025. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

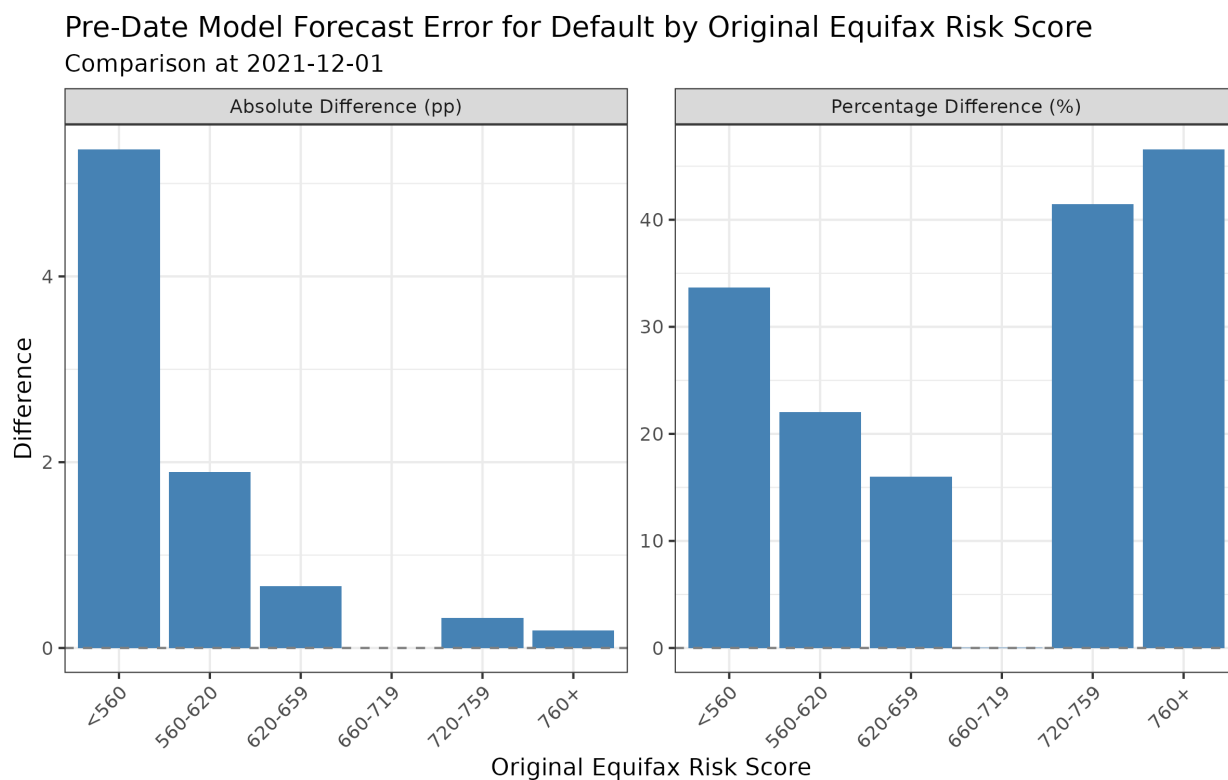


Figure 4.19: Cumulative 2-year forecast differences relative to actual outcomes, by original risk score groups. Cumulative forecasts over the period 2020-2021, differences computed as of December 2021. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

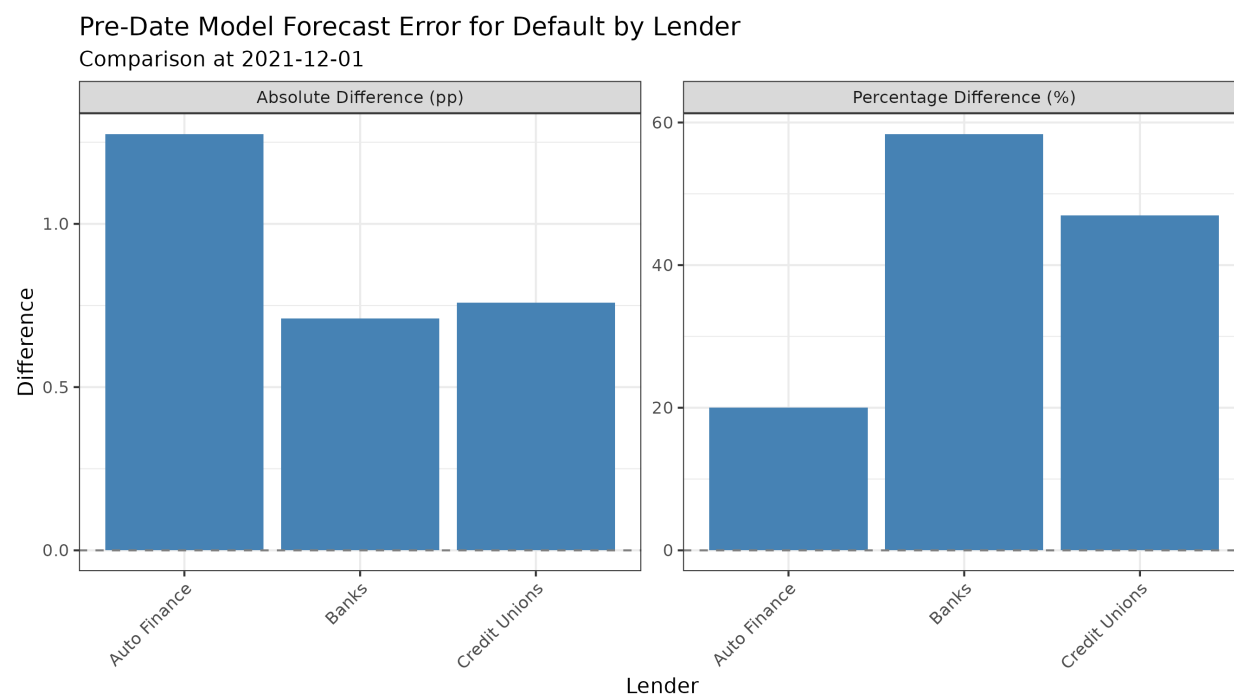


Figure 4.20: Cumulative 2-year forecast differences relative to actual outcomes, by lender. Cumulative forecasts over the period 2020-2021, differences computed as of December 2021. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

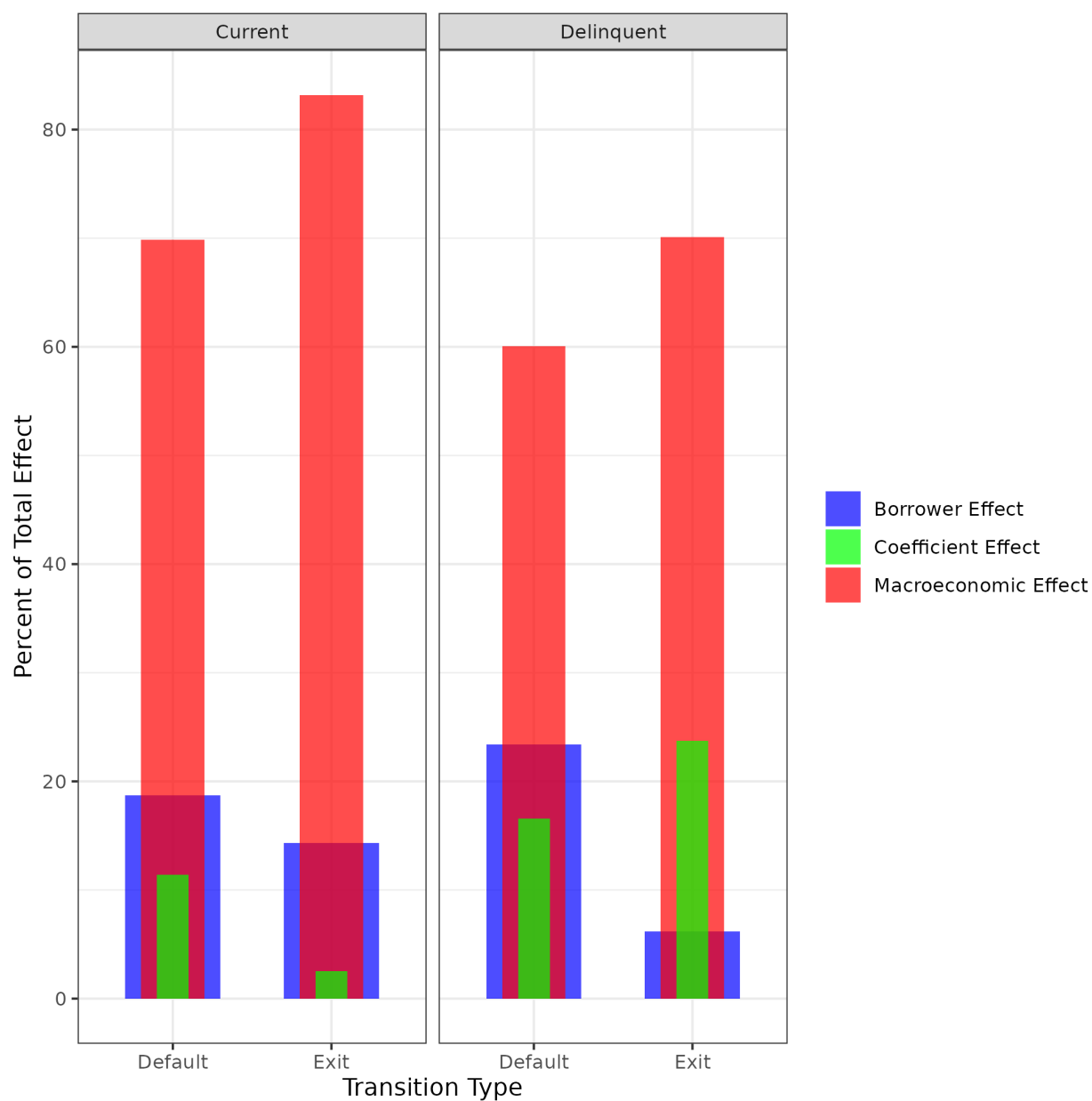


Figure 4.21: Oaxaca-Blinder decomposition on transition probabilities using the methods of Fairlie (2005). Baseline period is December 2019, alternative period is December 2021. Borrower and macro decompositions evaluated using alternative period coefficients. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

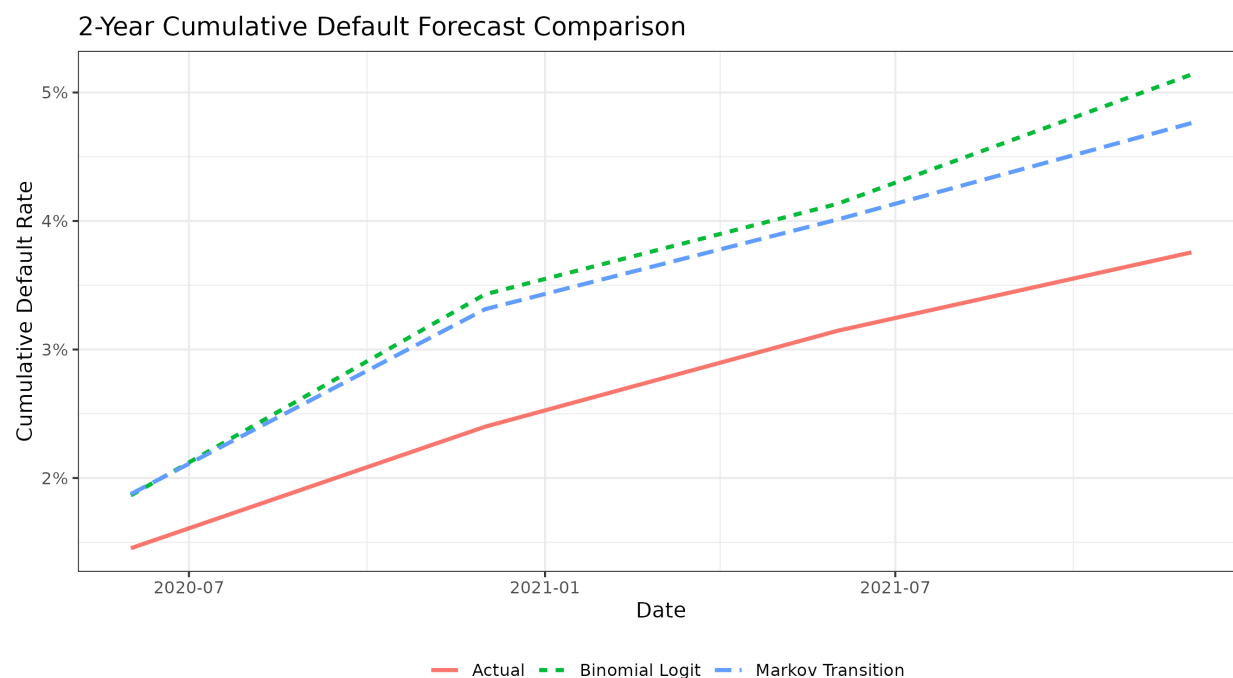


Figure 4.22: Comparison of Markov transition model and binomial logistic discrete-time hazard model cumulative 2-year forecasts from 2020-2021. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

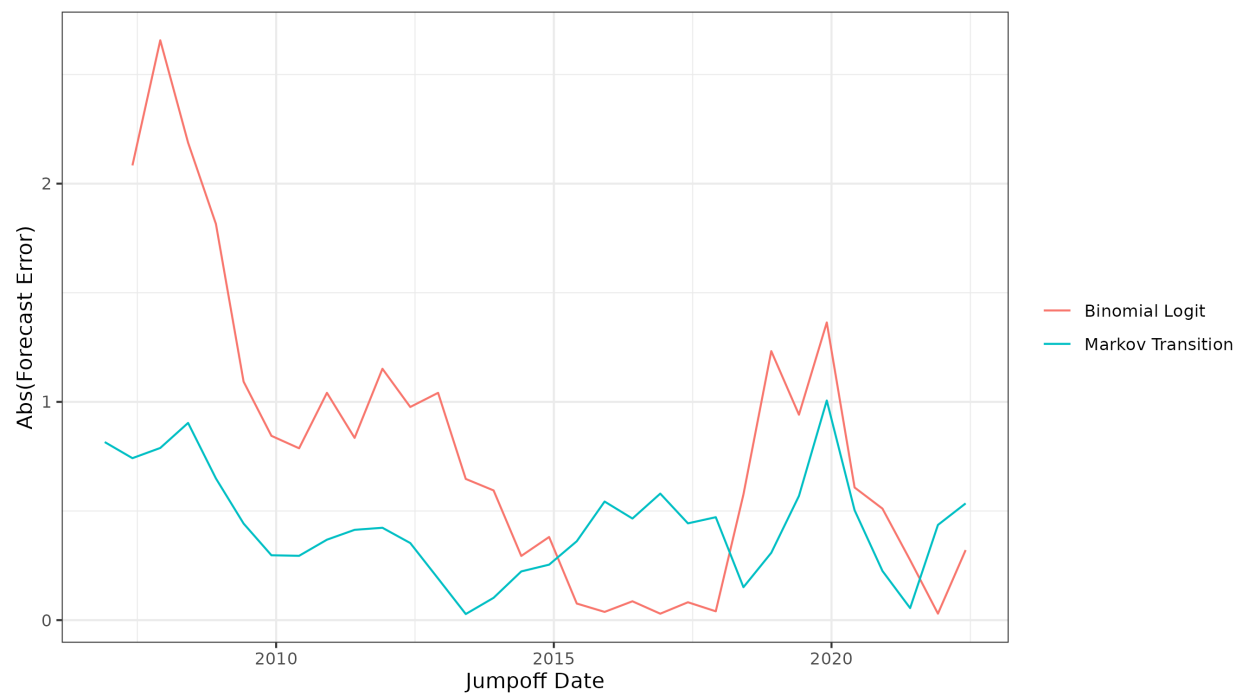


Figure 4.23: Comparison of rolling cumulative two-year default forecast errors between Markov transition model and binomial logistic discrete-time hazard model. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

4.12 Appendix

4.12.1 Hazard Specification

Prior to estimating our full model, we examine how loan age may explain transition probabilities so that we can specify the functional form to use for age in the main model. Different choices in modeling the hazard function (such as exponential, Gompertz, Weibull or non-parametric) embed assumptions about how risk evolves over time. An incorrectly specified hazard can lead to biased estimates of covariate effects, misleading predictions about survival probabilities, and erroneous conclusions about duration dependence. For instance, imposing a monotonically increasing hazard function when the true hazard is non-monotonic can mask critical periods of heightened or reduced risk. As Lancaster (1990) demonstrates, these specification errors can fundamentally alter conclusions derived from the estimated model.

We begin by estimating a model that conditions transitions only on various transformations of loan age and excludes other covariates. We estimate the model over 2000-2019 to remove effects from the pandemic period. Following Wu et al. (2018), we compare the following specifications: Exponential, Gompertz, quadratic, Weibull, and nonparametric. The exponential form assumes a constant hazard; thus, the loan age drops out of the model altogether. The Gompertz assumes a log-linear relationship between the log odds of default and the hazard, and the quadratic offers more flexibility by allowing for non-monotonicity. The Weibull specification results in a log transformation to be placed on the loan age. Finally, we specify a nonparametric hazard by creating indicator variables for each year of the loan's life, grouping all loans greater than 5 years old into one final category due to a relatively small number of observations in that grouping.

We assess the in-sample fit of these competing specifications so that the data inform our choice of a hazard function for the main model. Our ideal specification will balance model fit and predictive accuracy with a parsimonious specification to avoid over-fitting. We evaluate model specification along these criteria using the root mean squared error (RMSE),

McFadden's pseudo- R^2 , and the Bayesian information criteria (BIC). Given that much of our focus is on aggregate outcomes, we compute the RMSE as the difference between the average predicted probabilities and actual relative frequency of outcomes. The pseudo- R^2 and BIC are computed normally using all of the observations.

Tables 4.A1 and 4.A2 show the model selection statistics for the current and delinquent models. Across both models the quadratic and nonparametric approaches generally have the best results. For the current state, the nonparametric specification has the smallest RMSE, while the quadratic has the largest pseudo- R^2 and smallest BIC. In the delinquent model, the quadratic has the lowest RMSE and BIC, and a pseudo- R^2 comparable to the nonparametric specification. Given these results, we choose the quadratic specification in our analysis for its simplicity and parsimonious specification relative to the nonparametric approach.

4.12.2 Regression Results: 2000-2006

To examine estimated relationships prior to the GFC, we produce tables of our multinomial logit regression results estimated over the 2000-2006 period. Tables 4.A3 and 4.A4 show the results of our estimation for transitions from current and delinquent, respectively. As before, the coefficients represent the marginal effects on the relative risk of transition to a particular state compared with staying in the current state. Credit unions and origination risk scores greater than 760 are the reference groups for lender type and origination risk score.

Although the magnitudes of various parameter estimates differ in magnitude compared to the 2000-2019 sample period, the estimated relationships for lender type and loan characteristics that are statistically significant generally hold. For the macroeconomic variables, unemployment has a larger effect on the relative risk of transitioning from current-default in the 2000-2006 sample. Unemployment also has a counterintuitive effect on transitioning from both current-exit and delinquent-exit in the earlier sample, so that an increase in unemployment is associated with a higher relative risk of transition to that state. This could again be a result of measurement error from loans that were classified as exiting the sample

but that actually defaulted. The estimates for HPI generally agree with the longer sample period except for the current-default state, where there is a small but positive relationship between an increase in HPI and transitions from current-default. Across the macroeconomic variables, the absolute change in the unemployment coefficient for transitions from current-default between the 2000-2006 and 2000-2019 sample periods is relatively the largest.

4.12.3 Forecasts and Parameter Instability by Lender Type

We use the lender types associated with each loan in our data to examine if the relationship between unemployment and transitions to default and exit states changes differently in 2020 by lender type. To do this, we re-estimate the rolling out-of-sample framework on loans stratified by lender type so that we have lender-specific models and predictions. For statistical power, we focus on the broader lender types that we observe for our full sample period. Figures 4.A1, 4.A2, and 4.A3 show these results for auto finance, bank, and credit union loans, respectively.

Focusing on transitions to default, loans originated from auto finance companies generally have more frequent occurrences of default relative to other lender types. The absolute forecast error for transitions from current to default is lowest for the auto finance loans, suggesting the relationship between unemployment and default transitions may have been more stable. For banks and credit unions, the absolute forecast errors for transitions from current to default in 2020 are larger, both more than 1 percentage point higher than for auto finance loans. The other notable source of forecast error in 2020 is transitions from delinquent to exit. This discrete jump in predicted exits during 2020 occurs primarily for auto finance companies.¹⁴ We are cautious not to over-interpret those errors due to the potential for measurement error in that state.

Figures 4.A4, 4.A5, and 4.A6 confirm the coefficient instability on the unemployment variable is associated with the size of the forecast errors by lender. The largest amount of

¹⁴A discrete jump in the transitions from delinquent to exit also occurs for credit union loans, but not until 2021 when the underlying HPI data began to spike.

instability for transitions from current to default is observed for banks and credit unions. In both cases, the marginal effect of the year-over-year change in unemployment decreases by roughly 15 percentage points in 2020. After 2020, the estimated effect of unemployment is statistically zero for banks and credit unions, and remains that way until the end of the sample period. The rolling coefficients for unemployment on delinquent to exit transitions are statistically significant only for auto finance loans prior to 2020, and drop by roughly 10 percentage points to statistically zero levels after 2020. Despite some volatility in 2020, these coefficients are statistically zero for other lender types for most of the out-of-sample period.

4.12.4 Forecasts and Parameter Instability with Alternative Macroeconomic Specifications

Model without macroeconomic variables

Given the instability in the unemployment coefficients and the resulting forecast errors, it is instructive to understand how a model that excludes these variables might perform. We re-estimate the rolling regression model excluding macroeconomic variables, again treating June 2000-December 2006 as the initial in-sample period.

Figure 4.A7 plots the resulting semiannual out-of-sample forecasts. Not surprisingly, forecast errors for transition to default decline during the GFC and Covid recessions. However, forecast errors worsen for transitions to the exit state. This is reinforced by the two-year cumulative forecast from 2019-2021 using this model, as shown in Figure 4.A8. Relative to the model that includes macroeconomic variables, the cumulative forecast errors are smaller across models that use parameters from before or after 2020, but all of the parameter choices perform worse for the exit state. This suggests that a model without macro variables may improve forecasts for certain absorbing state transitions, but at the expense of others.

Model with additional macroeconomic variables

As an alternative to excluding macroeconomic variables, we consider additional macroeconomic variables beyond just year-over-year changes in unemployment and HPI to examine if the additional macroeconomic dynamics improve forecast accuracy. Following Heitfield and Sabarwal (2004), we include the annual difference in one-year Treasury rates to capture incentives to refinance during periods of low interest rates. While our data do not include information on loan-to-value ratios, we attempt to capture changes in the market value of auto loans by including the annual percentage change in the consumer price index for used vehicles obtained from the Bureau of Labor Statistics.¹⁵ We also control for the level of unemployment as opposed to only the year-over-year change in order to capture both observable deterioration and severity of unemployment at a given time. Figure 4.A9 plots the resulting semiannual out-of-sample forecasts. The forecasts are nearly identical to the forecasts that include only unemployment and HPI. This offers some evidence that the forecast errors and parameter instability are not a result of additional macroeconomic dynamics missing from the model.

4.13 Appendix Tables and Figures

¹⁵Specifically, we use the consumer price index for all urban consumers based on average prices of used cars and trucks in US cities.

Table 4.A1: Current State Model Hazard Fit Comparison.

Model	RMSE	Pseudo-R2	BIC
Exponential	17.3	0	1292080
Gompertz	6.2	0.048	1230154
Quadratic	7.4	0.051	1226555
Weibull	11.7	0.034	1248651
Nonparametric	6	0.048	1230430

Fit statistics for various specifications of the hazard function in the model of transitions from the current state. All specifications estimated only with a hazard function and no other covariates to determine best fit.

Table 4.A2: Delinquent State Model Hazard Fit Comparison.

Model	RMSE	Pseudo-R2	BIC
Exponential	12.5	0	76895
Gompertz	7.5	0.021	75304
Quadratic	6.5	0.025	75062
Weibull	8.8	0.019	75448
Nonparametric	7.1	0.025	75188

Fit statistics for various specifications of the hazard function in the model of transitions from the delinquent state. All specifications estimated only with a hazard function and no other covariates to determine best fit.

Table 4.A3: Current State Multinomial Logit Results (2000-2006)

	y = Transitions from Current State		
	Current-Delinquent	Current-Exit	Current-Default
Intercept	0.000*** [0.000, 0.000]	74.180*** [73.993, 74.368]	0.044*** [0.044, 0.044]
Lender: Auto Finance	1.487*** [1.425, 1.551]	0.905*** [0.880, 0.931]	1.290*** [1.253, 1.329]
Lender: Banks	1.098*** [1.061, 1.136]	0.951*** [0.923, 0.981]	0.748*** [0.733, 0.763]
Loan Age (months)	1.057*** [1.050, 1.064]	1.006*** [1.003, 1.008]	1.083*** [1.073, 1.093]
Loan Age ²	0.999*** [0.999, 1.000]	1.000*** [1.000, 1.000]	0.999*** [0.999, 0.999]
Equifax Risk Score: 720-759	2.377*** [2.371, 2.383]	1.015 [0.981, 1.050]	0.943*** [0.939, 0.947]
Equifax Risk Score: 660-719	6.693*** [6.602, 6.784]	1.022 [0.990, 1.055]	1.253*** [1.242, 1.265]
Equifax Risk Score: 620-659	16.983*** [16.323, 17.670]	1.054*** [1.016, 1.095]	1.975*** [1.960, 1.989]
Equifax Risk Score: <560	84.648*** [81.225, 88.215]	0.943*** [0.904, 0.984]	8.319*** [7.970, 8.683]
Borrower Age	0.993*** [0.990, 0.995]	0.996*** [0.995, 0.997]	0.991*** [0.988, 0.994]
First Mortgage Flag (t-1)	0.891*** [0.860, 0.924]	1.079*** [1.054, 1.105]	0.739*** [0.731, 0.746]
Ln(Balance) (t-1)	1.140*** [1.125, 1.155]	0.512*** [0.509, 0.515]	0.763*** [0.750, 0.776]
Joint Account	0.889*** [0.842, 0.940]	0.939*** [0.919, 0.960]	0.870*** [0.864, 0.876]
Annual Unemployment Change	1.020 [0.983, 1.059]	1.051*** [1.035, 1.067]	1.158*** [1.105, 1.213]
Annual HPI Change	0.987*** [0.982, 0.993]	1.009*** [1.007, 1.011]	0.993** [0.986, 1.000]
Seasonal Effect (Jan-June = 1)	0.828*** [0.783, 0.877]	1.044*** [1.022, 1.066]	0.814*** [0.810, 0.818]
Num.Obs.	305003		
R2	0.081		
R2 Adj.	0.081		
Pseudo-R2	0.08		
AUC	0.69		

Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations. Model estimated on sample data from 2000-2006. Coefficients represent relative risks. *, **, *** represent statistical significance at 0.1, 0.05, and 0.01 levels, respectively. 95 percent confidence intervals shown in brackets. Standard errors are clustered at the tradeline level.

Table 4.A4: Delinquent State Multinomial Logit Results (2000-2006)

	y = Transitions from Delinquent State		
	Delinquent-Delinquent	Delinquent-Exit	Delinquent-Default
Intercept	0.277*** [0.276, 0.278]	17.602*** [17.509, 17.695]	19.453*** [19.349, 19.556]
Lender: Auto Finance	1.221*** [1.104, 1.349]	0.847*** [0.753, 0.954]	0.938 [0.829, 1.061]
Lender: Banks	1.010 [0.933, 1.094]	1.115** [1.004, 1.239]	0.768*** [0.701, 0.842]
Loan Age (months)	0.982* [0.963, 1.002]	0.985 [0.960, 1.011]	0.951*** [0.930, 0.973]
Loan Age ²	1.000*** [1.000, 1.001]	1.001*** [1.000, 1.001]	1.001*** [1.000, 1.001]
Equifax Risk Score: 720-759	1.428*** [1.414, 1.442]	0.522*** [0.518, 0.526]	2.044*** [2.023, 2.066]
Equifax Risk Score: 660-719	1.151* [0.983, 1.349]	0.535*** [0.452, 0.633]	1.428*** [1.235, 1.651]
Equifax Risk Score: 620-659	1.759*** [1.526, 2.028]	0.652*** [0.546, 0.779]	1.695*** [1.427, 2.014]
Equifax Risk Score: <560	2.306*** [2.075, 2.564]	0.683*** [0.594, 0.785]	3.173*** [2.798, 3.598]
Borrower Age	1.001 [0.996, 1.006]	0.999 [0.992, 1.006]	0.989*** [0.982, 0.995]
First Mortgage Flag (t-1)	0.850** [0.726, 0.996]	0.952 [0.769, 1.178]	0.647*** [0.526, 0.796]
Ln(Balance) (t-1)	1.039* [0.998, 1.081]	0.671*** [0.635, 0.709]	0.726*** [0.694, 0.759]
Joint Account	0.932 [0.820, 1.059]	0.953 [0.799, 1.137]	0.780*** [0.669, 0.910]
Annual Unemployment Change	1.007 [0.922, 1.099]	1.061 [0.939, 1.200]	0.989 [0.893, 1.096]
Annual HPI Change	0.996 [0.985, 1.008]	0.994 [0.979, 1.010]	0.981*** [0.967, 0.994]
Seasonal Effect (Jan-June = 1)	0.886* [0.783, 1.003]	0.941 [0.793, 1.117]	0.801*** [0.693, 0.927]
Num.Obs.	5992		
R2	0.030		
R2 Adj.	0.030		
Pseudo-R2	0.03		
AUC	0.61		

Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations. Model estimated on sample data from 2000-2006. Coefficients represent relative risks. *, **, *** represent statistical significance at 0.1, 0.05, and 0.01 levels, respectively. 95 percent confidence intervals shown in brackets. Standard errors are clustered at the tradeline level.

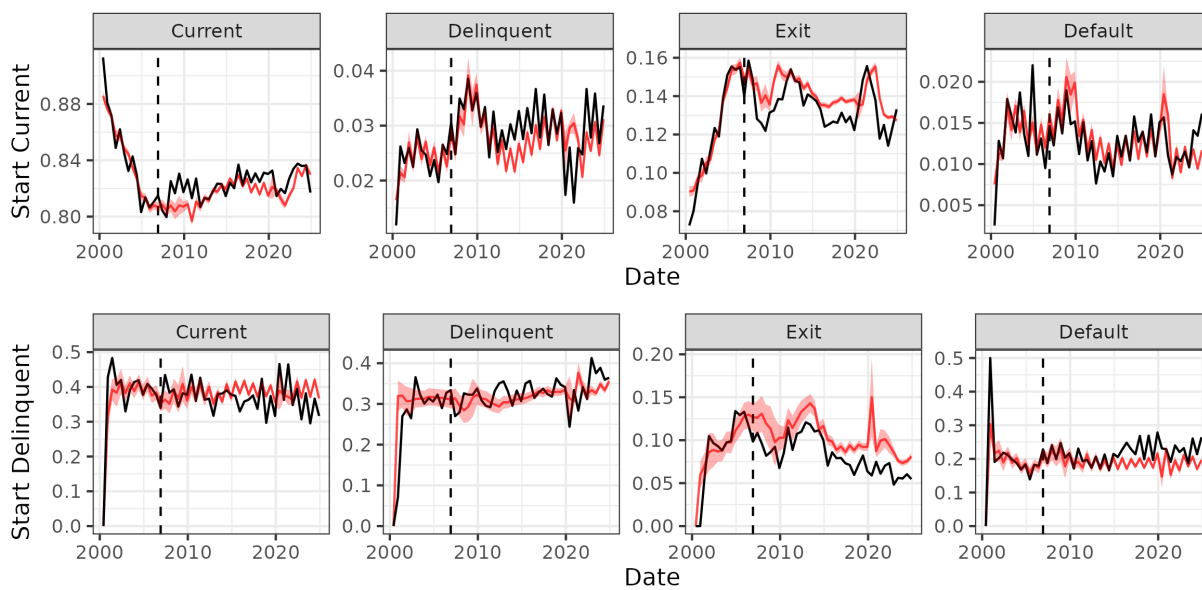


Figure 4.A1: Auto finance loans: Time path of current and delinquent-state transition probabilities from pseudo out-of-sample estimation. Estimation variables include loan age, lender type, loan characteristics, and year-over-year changes in US state unemployment and HPI. The red line represents pseudo out-of-sample forecasts and the black line represents actual outcomes. Red shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Dashed line indicates final in-sample forecast as of December 2006. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

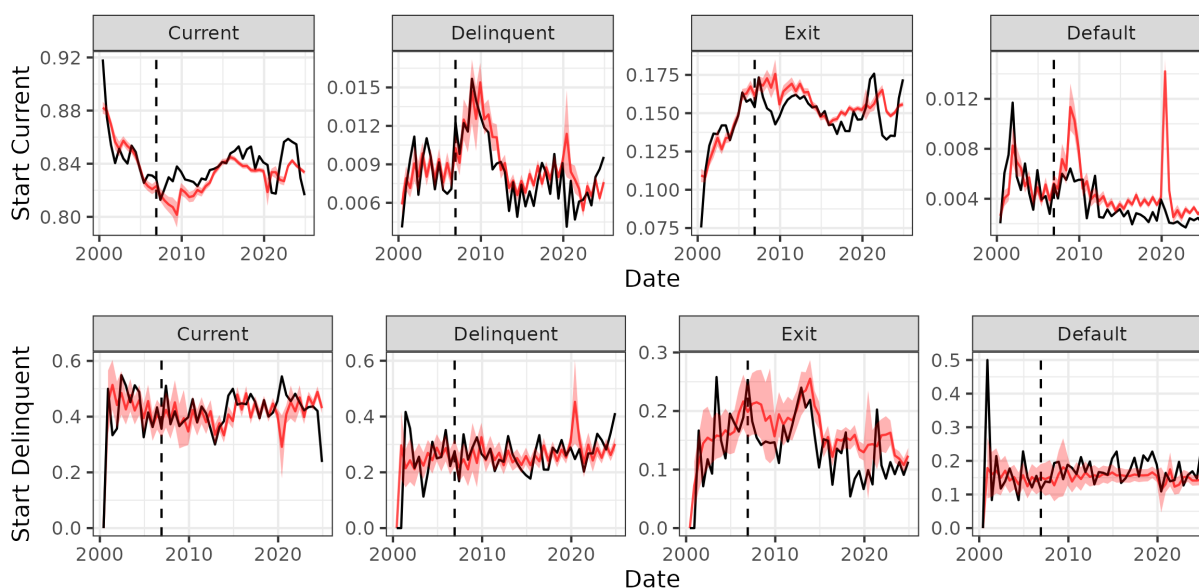


Figure 4.A2: Bank loans: Time path of current and delinquent-state transition probabilities from pseudo out-of-sample estimation. Estimation variables include loan age, lender type, loan characteristics, and year-over-year changes in US state unemployment and HPI. The red line represents pseudo out-of-sample forecasts and the black line represents actual outcomes. Red shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Dashed line indicates final in-sample forecast as of December 2006. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

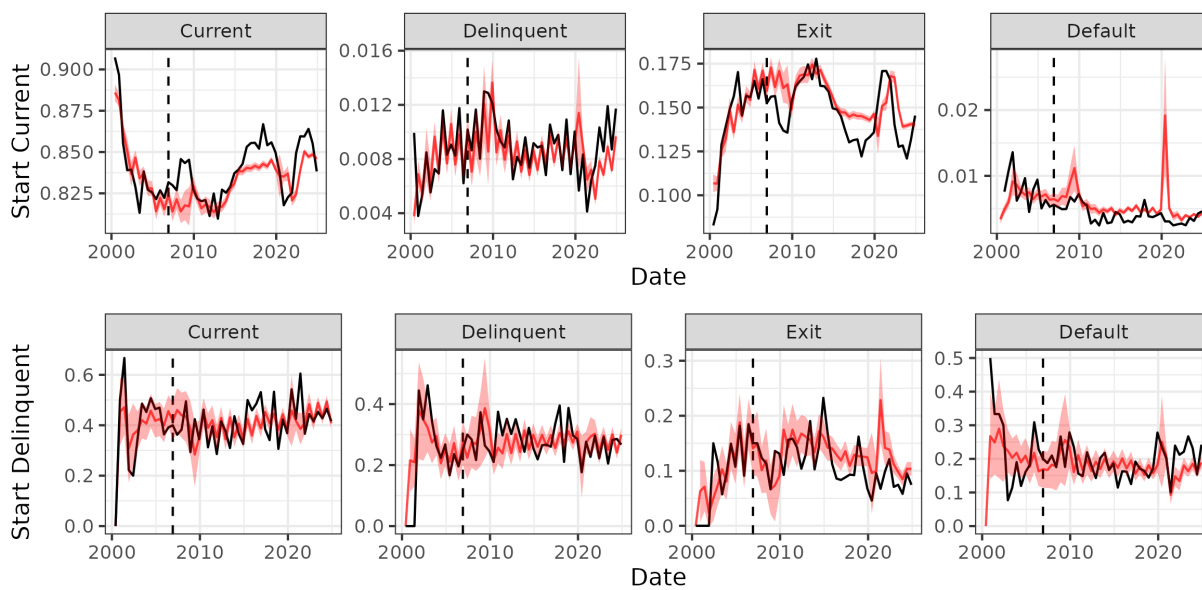


Figure 4.A3: Credit union loans: Time path of current and delinquent-state transition probabilities from pseudo out-of-sample estimation. Estimation variables include loan age, lender type, loan characteristics, and year-over-year changes in US state unemployment and HPI. The red line represents pseudo out-of-sample forecasts and the black line represents actual outcomes. Red shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Dashed line indicates final in-sample forecast as of December 2006. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

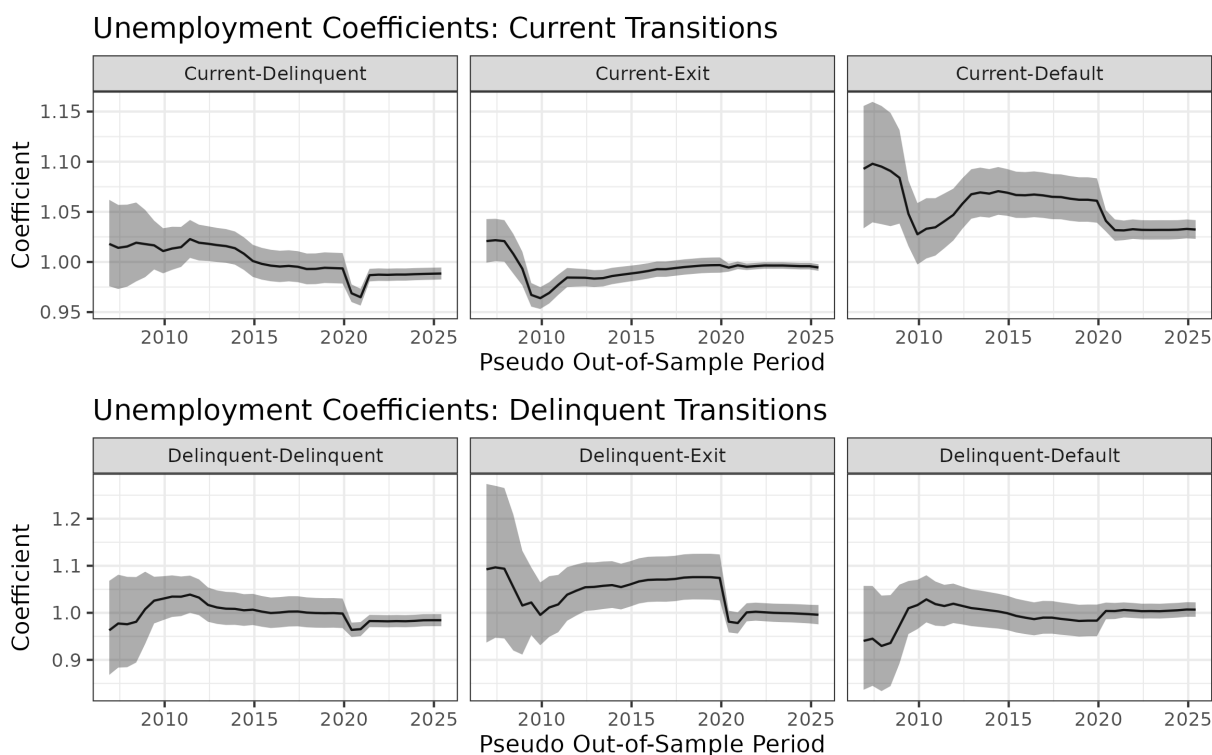


Figure 4.A4: Auto finance loans: Multinomial logit estimates of US state unemployment coefficients during pseudo out-of-sample estimations, 2006-2025. Estimation variables include loan age, lender type, loan characteristics, and year-over-year changes in US state unemployment and HPI. Shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

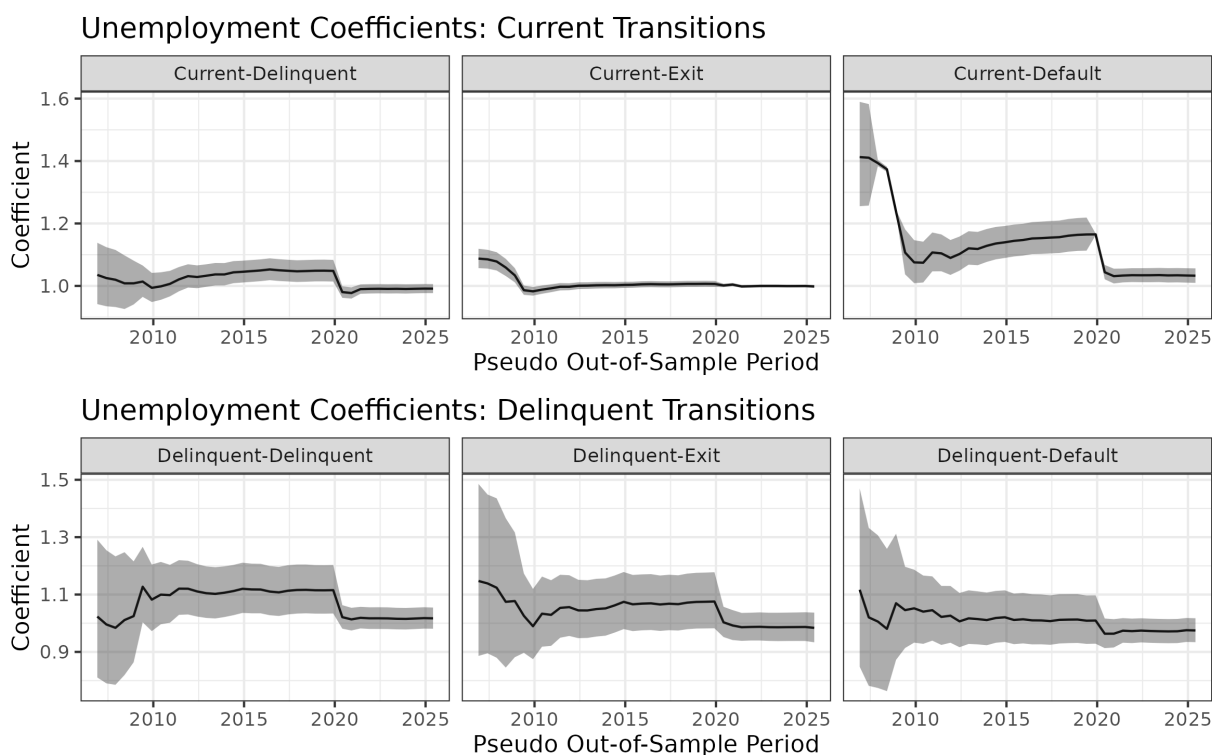


Figure 4.A5: Bank loans: Multinomial logit estimates of US state unemployment coefficients during pseudo out-of-sample estimations by model, 2006-2025. Estimation variables include loan age, lender type, loan characteristics, and year-over-year changes in US state unemployment and HPI. Shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

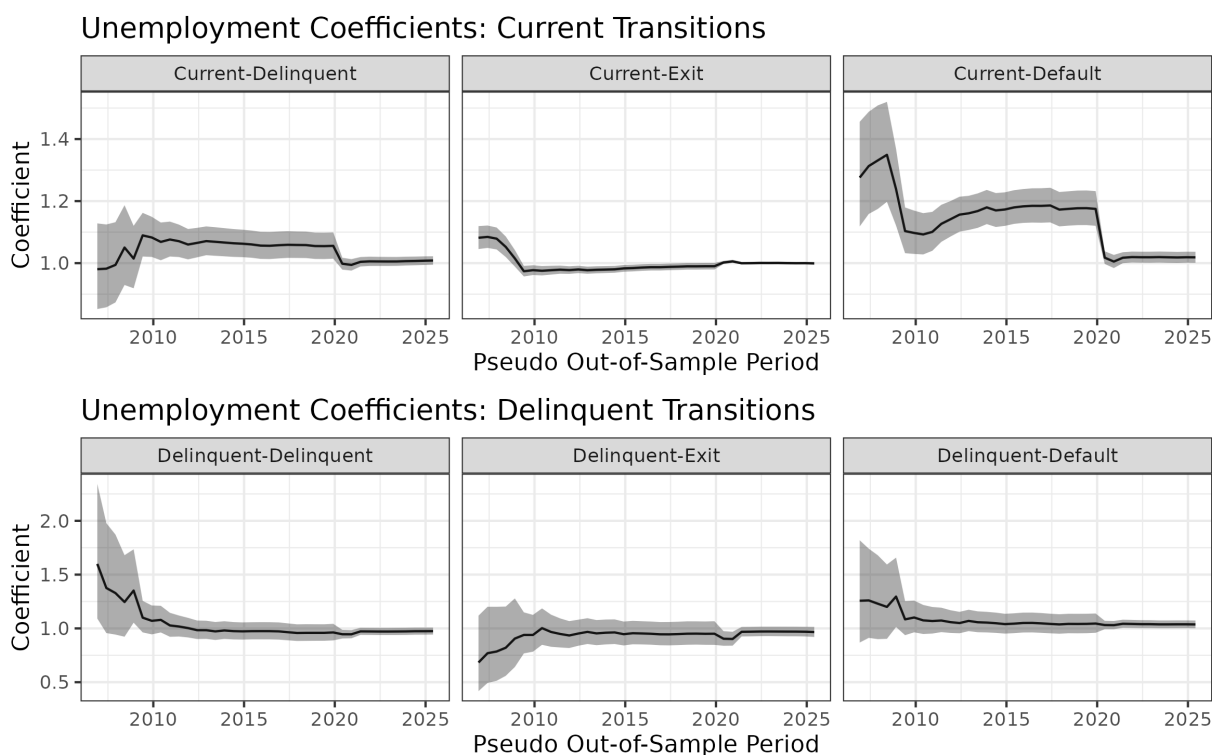


Figure 4.A6: Credit union loans: Multinomial logit estimates of US state unemployment coefficients during pseudo out-of-sample estimations, 2006-2025. Estimation variables include loan age, lender type, loan characteristics, and year-over-year changes in US state unemployment and HPI. Shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

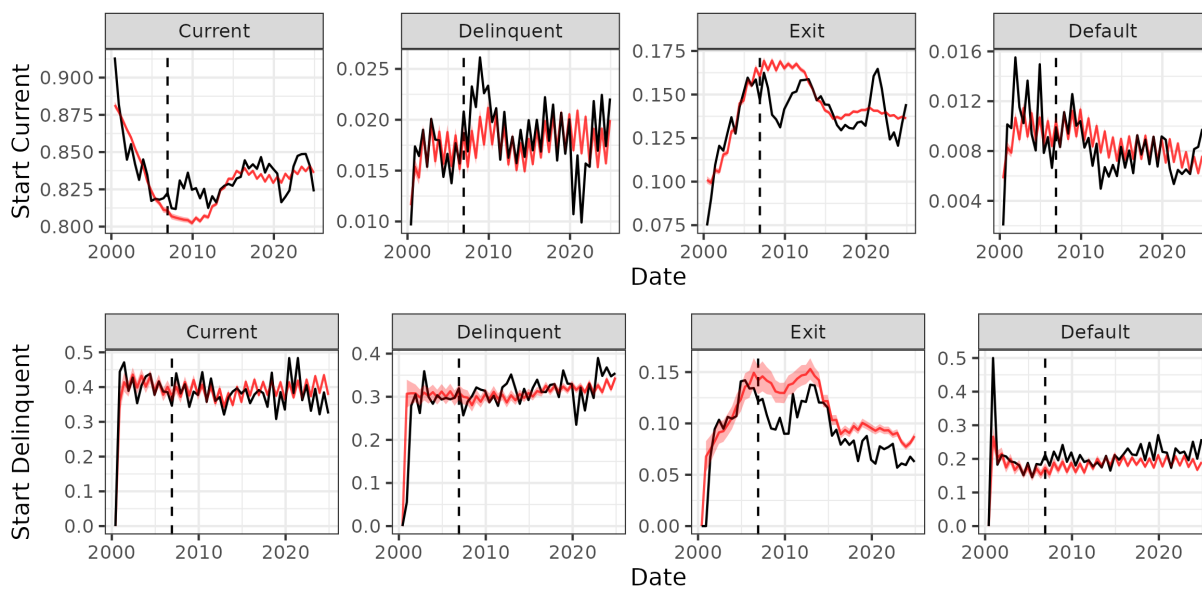


Figure 4.A7: Model without macroeconomic variables: Time path of current and delinquent-state transition probabilities from pseudo out-of-sample estimation. Estimation variables include loan age, lender type, loan characteristics, but excludes any macroeconomic variables. The red line represents pseudo out-of-sample forecasts and the black line represents actual outcomes. Red shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Dashed line indicates final in-sample forecast as of December 2006. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

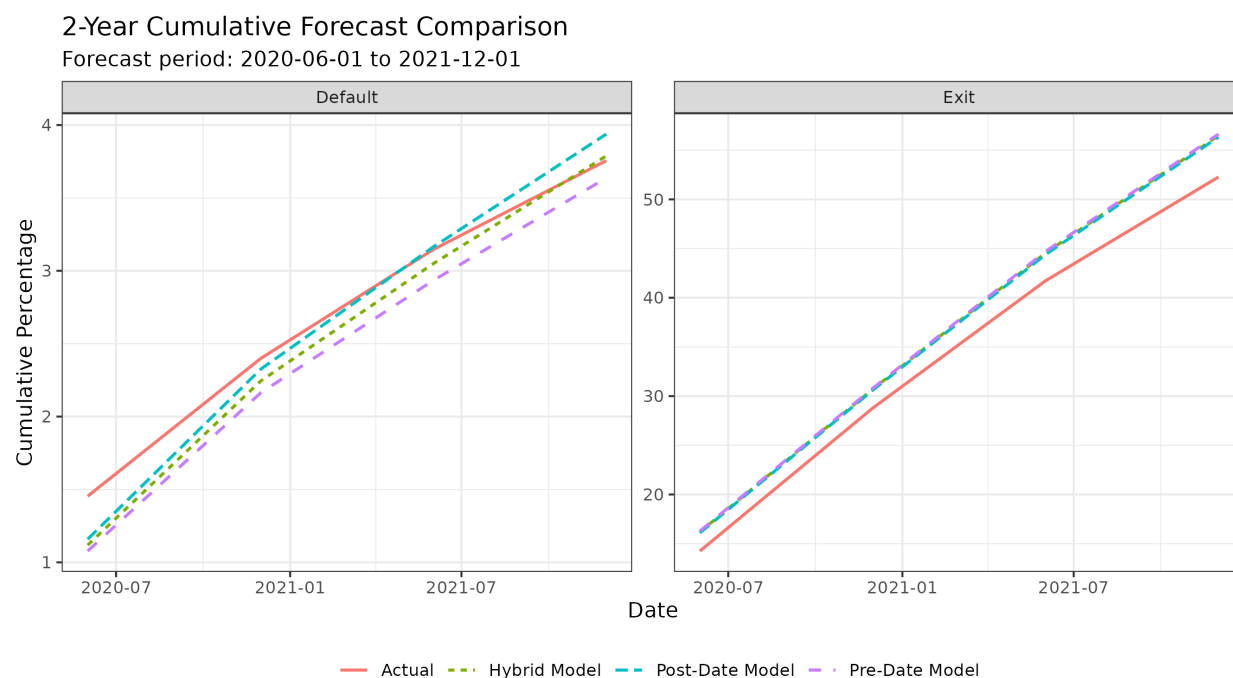


Figure 4.A8: Model without macroeconomic variables: Cumulative 2-year forecast differences relative to actual outcomes. Cumulative forecasts over the period 2020-2021, differences computed as of December 2021. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

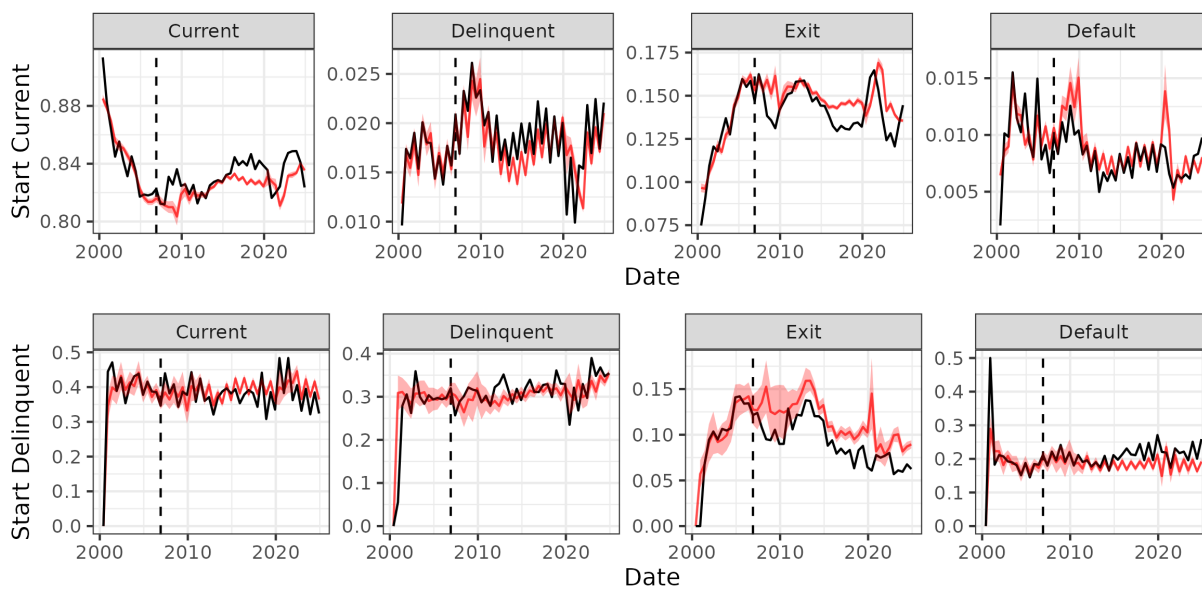


Figure 4.A9: Model without additional variables: Time path of current and delinquent-state transition probabilities from pseudo out-of-sample estimation. Loan-level estimation variables include loan age, lender type, loan characteristics. Macroeconomic variables include the level of unemployment, the absolute year-over-year change in US state unemployment, the 1-year Treasury rate, the annual percentage change in HPI, and the consumer price index for used vehicles. The red line represents pseudo out-of-sample forecasts and the black line represents actual outcomes. Red shaded areas represent 95 percent confidence intervals, where standard errors are clustered at the loan level. Dashed line indicates final in-sample forecast as of December 2006. Source: FRBNY Consumer Credit Panel / Equifax Data, Authors' calculations

Chapter 5

Probabilistic Anchoring of Inflation Expectations

5.1 Introduction

Economists agree that “anchored” inflation expectations are a *sine qua non* for successful and credible monetary policy. This consensus reflects both theoretical insights and practical experience. In workhorse New Keynesian models, anchored expectations mean that temporary shocks to the economy will have less persistent effects on both actual inflation and real economic activity (Woodford, [2003](#); Galí, [2008](#)). When households and firms believe inflation will return to target following a shock, they are less likely to adjust wages and prices in ways that propagate the shock. Empirically, central banks with credible inflation targets have achieved both lower average inflation and less output volatility—the “Great Moderation” prior to 2007 is often attributed in part to better-anchored expectations (Bernanke, [2004](#); Clarida et al., [2000](#)). In the United States, the FOMC formalized its inflation target at 2 percent in January 2012, and maintaining well-anchored expectations around this target has been a central policy objective since.

While the FOMC’s stated target is 2 percent, central banks universally recognize that

inflation will not remain at exactly this level at all times. Rather, policymakers tolerate deviations within some acceptable range. The FOMC most recently formalized this tolerance in its June 2025 “Statement on Longer Run Goals,” stating that: “[in] order to anchor longer-term inflation expectations at this level [2 percent], the Committee seeks to achieve inflation that averages 2 percent over time, and therefore judges that, following periods when inflation has been running persistently below 2 percent, appropriate monetary policy will likely aim to achieve inflation moderately above 2 percent for some time.” Other central banks similarly tolerate deviations within some range deemed *acceptable*. The Bank of Canada formalizes a “control range” of 1–3 percent, the ECB communicates its inflation target as “close to 2 percent,” and if they miss their 2 percent target by more than 1 percent on either side, the Bank of England Governor is obliged to publish a letter explaining why.

These statements and monetary policy frameworks, in their different ways, all recognize the inherent uncertainty in inflation dynamics, the recurrence of shocks, and the practical limits of monetary policy control given “long and variable lags.” Indeed, standard New Keynesian models show that even under optimal monetary policy, inflation deviates from target with variance proportional to the variance of (unavoidable) cost-push and other real shocks (cf. Galí (2008), equations (6)-(7), page 99). The optimal variance reflects the policymaker’s tradeoff between inflation and output stabilization: eliminating all inflation volatility would require unacceptably large output fluctuations. This provides a theoretical foundation for specifying the policymaker’s target density for inflation as a distribution centered at 2% rather than a point target. The question of whether inflation expectations are anchored therefore involves not just whether agents’ mean expectations equal 2 percent, but whether their distribution of beliefs is consistent with the policymaker’s tolerance for deviations. As Reis (2021) emphasizes, a well-anchored expectations distribution should have most probability mass concentrated near the target, with limited probability assigned to extreme outcomes—what he terms “inflation disasters.”

Despite this unanimity on the importance of having anchored expectations, there is less

agreement about how to measure whether inflation expectations are anchored; see Bems et al. (2021) and Kumar et al. (2015). The metrics proposed in the literature can be broadly viewed as falling into two camps, what Cline and Rich (2026) call “level” and “shock” anchoring.¹ First, expectations are deemed level-anchored when economic agents’ mean (or median) expectations are aligned with the inflation target—that is, when deviations from 2 percent are “small,” ideally mean zero when the policymaker’s loss function is quadratic.² Second, shock-anchoring means that long-run inflation expectations should not respond to short-run shocks. That is, agents should not revise their long-run inflation expectations with incoming macroeconomic news. Bernanke (2007) captured this second perspective in a speech: “I use the term ‘anchored’ to mean relatively insensitive to incoming data. So, for example, if the public experiences a spell of inflation higher than their long-run expectation, but their long-run expectation of inflation changes little as a result, then inflation expectations are well anchored.” Following this logic, papers such as Jochmann et al. (2010), Nautz and Strohsal (2015), and Kumar et al. (2015) test for anchoring by examining whether long-run inflation expectations respond to short-run expectations or macroeconomic surprises.

An important limitation of this literature is its almost exclusive focus on point expectations—the mean or median of agents’ inflation forecasts. Yet economic agents form subjective density forecasts for future inflation that capture not just their central expectation, but also the risks and uncertainties they perceive (Boero et al., 2008; Clements, 2014). These underlying density forecasts may be tight or dispersed, symmetric or skewed. Moreover, they may well change over time, for example if inflation uncertainty is not constant over time. From a risk-management perspective, it is the probability of extreme inflation realizations that may matter most for welfare and policy (Reis, 2021). It is perfectly possible that point

¹ An important third metric used to assess anchoring is to examine individual-level expectations of inflation and assess their dispersion. Inflation anchoring is taken to require low dispersion or disagreement across different agents’ expectations; e.g., see Capistrán and Ramos-Francia (2010), Kumar et al. (2015), and Naggert et al. (2023). Our empirical focus in this paper, however, is testing probabilistic anchoring of market or aggregated expectations.

² Under an asymmetric loss function in inflation deviations (errors), it can be optimal for such deviations to be biased; e.g., see Capistrán (2008).

inflation expectations appear well-anchored according to existing metrics, while this masks probabilistic unanchoring. For example, the mean expectation might remain near 2 percent, but the probability mass in the tails could increase substantially—agents might assign high probability to “inflation disasters” (persistent high inflation) or deflation.³ Alternatively, incoming macroeconomic news might not move the mean expectation, but could affect the tails or skewness of the distribution, thereby changing the perceived probability of extreme outcomes.

Our contribution. In this paper, we propose a generalization of existing measures of anchoring that extends the literature in two main directions. First, we unify the two traditional camps—testing for unbiasedness and insensitivity to news—into a single framework that reinterprets both as a Mincer-Zarnowitz-type efficiency regression (cf. Mincer and Zarnowitz (1969)), but applied to densities rather than point forecasts. A key conceptual contribution is formalizing what we call the policymaker’s *target density*—the complete probability distribution that reflects not just the 2 percent inflation target but also the policymaker’s tolerance for deviations around it. Inflation expectations are then deemed probabilistically anchored when, on average over a given sample period, the entire density of expectations is both unbiased—equal to this target density $\mu^*(s)$ —and orthogonal to information known when the expectation was formed.

Second, and more substantively, we extend measures of inflation anchoring designed for point expectations to propose a probabilistic analogue of anchored inflation expectations. Reflecting the growing availability of density forecast data from both surveys and financial markets (Kitsul and Wright, 2013; Grishchenko et al., 2019), we define inflation expectations to be probabilistically anchored when two conditions hold:

1. Unconditional anchoring: Agents’ density expectations of long-run inflation should be consistent with the policymaker’s target density, $\mu^*(s)$, which embeds both the

³As already intimated, there can also, when agents’ underlying subjective density functions for future inflation are asymmetric and/or when their loss function is not quadratic, be important differences between the mean inflation expectation and the “optimal” point expectation; e.g., see Gneiting (2011).

target level of inflation (2 percent) and the policymaker’s perceived degree of tolerance for deviations around this target. This generalizes the traditional condition that the mean inflation expectation equals the target level – $E[\pi_t^e] = \pi^* = 2\%$ – by requiring that the entire distribution of inflation expectations equals the policymaker’s target distribution.

2. Conditional anchoring: Long-run density expectations should be unrelated to incoming macroeconomic surprises across the full support of the expectations distribution. This extends Bernanke (2007)’s condition that mean expectations should be insensitive to news, to require that every point on the density—including the tails—should be insensitive to news. We show how functional coefficients reveal not just whether macro surprises affect expectations, but where on the density they matter. Do macro surprises move the center or the tails of the inflation expectations density?

Our framework nests and generalizes several existing concepts that, like us, accommodate some form of deviation of inflation expectations from target. Kumar et al. (2015)’s notion of “ ε -anchored” expectations asks whether the mean expectation falls within a tolerance band $[\pi^* - \varepsilon, \pi^* + \varepsilon]$. Our “regional probabilistic anchoring” asks instead whether the expected probability mass within this band matches the target probability. Similarly, Grishchenko et al. (2019) propose measuring anchoring through the probability of inflation falling within specific ranges around the target. We extend their approach by testing whether movements in these probabilities can be explained by macroeconomic surprises and by examining the full distributional response rather than selected regions or discrete event-probabilities. Kitsul and Wright (2013) also focus on specific regions, examining whether their RND-based probabilities of deflation or inflation exceeding 4% respond to macroeconomic surprises, though they do not explicitly frame this as a test of anchoring.

We develop methods to test probabilistic anchoring using function-on-scalar regression techniques borrowed from the functional data analysis (FDA) literature. In so-doing, we follow a small but growing literature in economics and econometrics (Meeks and Monti,

2023; Chang et al., 2024; Clements and Wang, 2024; Huber et al., 2024a) in using functional regression methods. Our contribution econometrically is to focus on inference and interpretation of the functional regression coefficients – so that we can test probabilistic anchoring. Our dependent variable is a time-varying probability density function—specifically, the risk-neutral density of long-run inflation implied by inflation options markets. Because probability densities are compositional data (constrained to be non-negative and integrate to one), linear operations require careful treatment. We work in the Bayes space $\mathcal{B}^2$ using the centered log-ratio (CLR) transformation (Van den Boogaart et al., 2014; Hron et al., 2016; Talská et al., 2018), which allows us to perform standard regression while ensuring our predictions remain valid densities. The functional coefficients reveal not just whether macro surprises affect expectations (as in point regressions), but where on the density they matter – center, tails, or specific inflation levels.

Following the application in Kitsul and Wright (2013), we apply our new anchoring framework to weekly options-implied inflation probability density forecasts extracted from inflation caps and floors from 2012 through 2025. These inflation derivatives offer information about the complete expected probability density for inflation five-years-ahead, including tail risks and uncertainty. They allow us to assess not just whether longer-run mean expectations are anchored, but whether the full distribution—including tail probabilities—aligns with the policymaker’s target. An advantage of using market-based estimates of inflation probability density functions is that they are directly observable at high frequency from traded securities, providing incentive-compatible measures of beliefs (plus risk premia). They allow precise identification of how expectations respond to macroeconomic surprises. In contrast, survey-based measures of households’ or firms’ subjective probability distributions of future inflation are typically available at a lower frequency and involve no financial consequences and may reflect strategic considerations.⁴

⁴An ideal approach would combine our analysis of market-based densities with parallel analysis of survey-based distributions where available (e.g., SPF density forecasts). We leave this complementary analysis to future work, noting that survey-based densities are often available only at quarterly frequency and with coarse binning, limiting the power to detect the tail movements we document at weekly frequency.

The plan of the remainder of this paper is as follows. Section 2 presents our data on risk-neutral densities and macroeconomic surprises. Section 3 sets out our conceptual framework, defining probabilistic anchoring and its relationship to existing concepts of anchoring. Section 4 develops the econometric methodology, explaining the Bayes space framework, CLR transformation, and function-on-scalar regression. Section 5 presents our empirical results, including rolling regressions, functional coefficients, and regional tests for anchoring. Section 6 concludes.

5.2 Data

5.2.1 Option-Implied Risk-Neutral Inflation Densities

Risk-neutral densities (RND) provide information on the distribution of investors' expectation of market prices under risk-neutral preferences. Breeden and Litzenberger (1978) show that the entire RND of the future value of an asset can be extracted using options traded on the underlying asset. They show that, under no-arbitrage conditions and a dense set of traded options with evenly-spaced strike prices and the same expiration date, the second derivative of the call option value with respect to the exercise price approximates the probability density function of the future asset price. Various empirical techniques exist to extract RNDs from the option data based on these results.

A key limitation when interpreting the results presented throughout this paper is that risk-neutral densities (RNDs, or Q-distributions) differ systematically from real-world probability distributions (P-distributions). RNDs would equal real-world densities only if all investors were risk-neutral, meaning asset prices would reflect expected future payoffs discounted at the risk-free rate. In reality, however, investors are typically risk-averse and demand compensation for bearing risk. Consequently, asset prices incorporate risk premia in addition to risk-free discounting. This risk premium, the difference between the required return on a risky asset and the risk-free rate, is what drives the wedge between the Q-

distribution and the P-distribution. Given this, our results do not disaggregate changes in real world probabilities versus changes in risk premia over time.

Despite this limitation, RNDs are an appropriate object of study for several reasons. First, from a policy perspective, central banks care about market- based measures because they affect asset prices, financing costs, and the transmission of monetary policy—regardless of whether movements reflect belief changes or risk premia shifts. If markets price elevated tail probabilities of high inflation, this matters for bond yields, corporate investment decisions, and financial stability whether due to increased subjective probabilities or higher risk compensation. Second, elevated inflation risk premia themselves signal concerns about inflation outcomes severe enough that investors demand compensation, indicating a form of distributional unanchoring in market confidence even if subjective point expectations remain stable.

We use weekly 5-year ahead inflation option-implied RND data over the period from July 2012 to May 2025. Due to data limitations on cap and floor data, we collect data that summarize estimated RNDs made available by the Federal Reserve Bank of Minneapolis (FRB Minneapolis) rather than construct it ourselves. These densities are constructed using US inflation caps and floors collected from Bloomberg that have the US consumer price index (CPI) as the underlying index. While granular data on the fitted densities are not made public, FRB Minneapolis makes publicly available selected quantiles and the first four moments of the weekly estimated RNDs. Each set of parameters are mapped to a particular date and observed at a weekly frequency. Documentation on construction of the RNDs note that the method used is a variant of Kitsul and Wright (2013), who use local polynomial regression to construct their RNDs.⁵

We use the moments provided by FRB Minneapolis to reconstruct continuous density functions. Given that nonparametric regression methods may be used to construct the

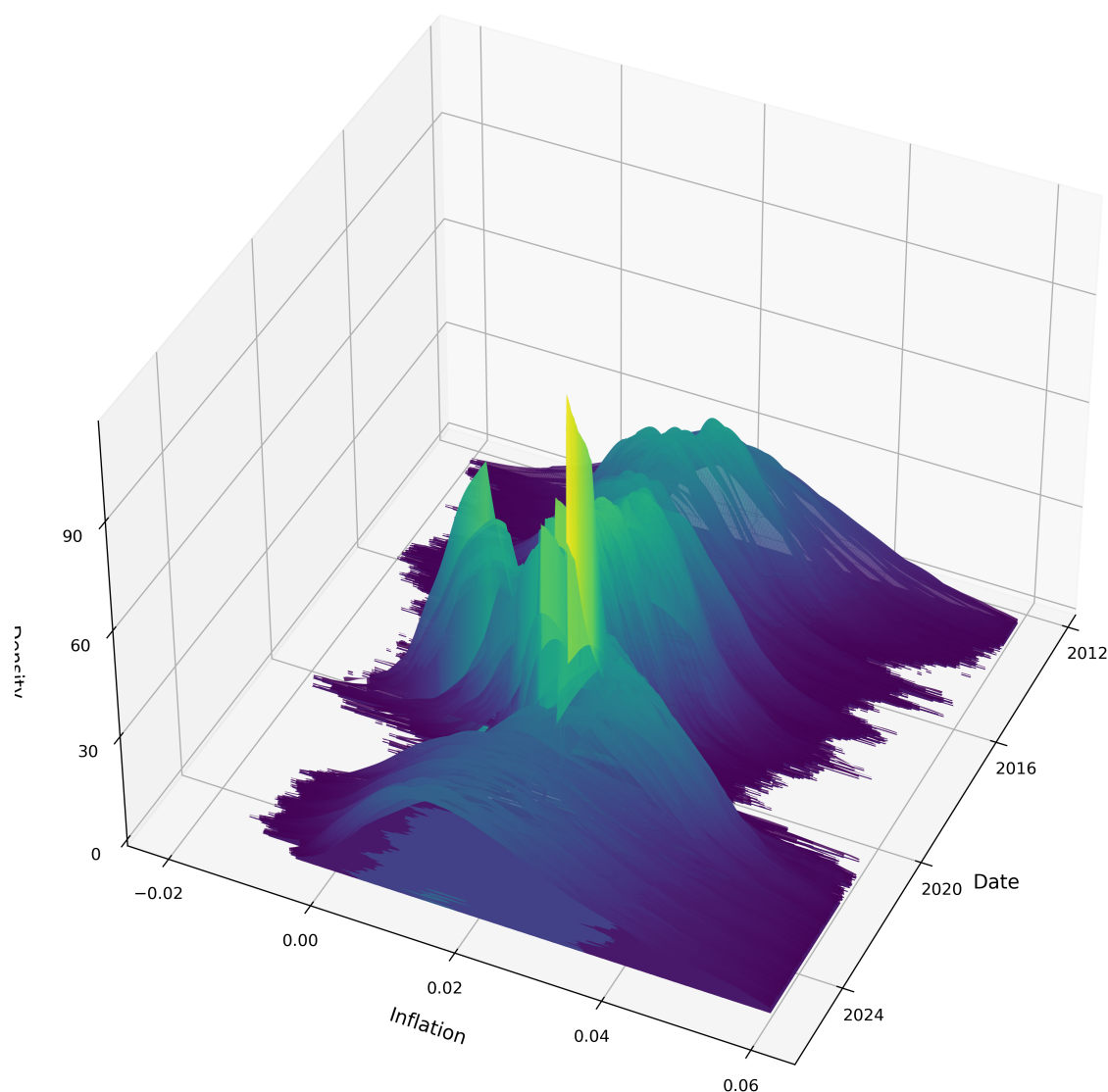
⁵See the RND data and additional documentation on the FRB Minneapolis website for the methodology used to construct the densities here: <https://www.minneapolisfed.org/banking/current-and-historical-market--based-probabilities>.

densities, we reconstruct continuous density functions using the published moments of the RNDs fitted to a Pearson distribution. The Pearson distribution takes the mean, standard deviation, skewness, and kurtosis as its parameters, and offers a relatively flexible approach to constructing the RNDs while nesting potentially simpler distributional forms that may be suggested by the parameters. We create a matrix containing a set of discrete density vectors over a common support grid, where each row is a density for a given time period, and each column is a discrete support point. Our support grid spans the space implied by available market strike prices on the US inflation cap and floor data, ranging from -2 percent inflation to 6 percent inflation. We discretize the support grid in increments of 1 basis point.

Some data cleaning on the RNDs is required. First, while our data begin in July 2012, FRB Minneapolis makes available 5-year ahead inflation RND data going back to 2009. We wish to maintain a weekly frequency of RND data, and in any cases where a week's data is missing we hold the previous week's density constant. There are a relatively large number of missing weeks between 2009 and July 2012, and only a small number of missing weeks after, so we drop observations before July 2012 in order to make fewer assumptions on the data. Second, there is one density on July 28, 2021 that is extremely concentrated in its center mass with very narrow tails. This is due to the kurtosis parameter obtained in the underlying FRB Minneapolis data increasing to a relatively large value on that date. To deal with this, we winsorize the density data at the top 2 percent of the data and renormalize the densities to integrate to one. This choice is primarily cosmetic for visualization purposes, and our main results are not affected by this choice relative to other options such as leaving the outlier density as it is or removing that density altogether.

Figure 5.1 plots the surface implied by our RND distributions. We can begin to see considerable variation in location, scale, and skewness over time. Specifically, we observe a rightward location drift of the densities from 2020-2021 and a consistently long right tail appearing in the distributions afterward. We also see considerable volatility, with occasional spikes, in the inflation densities through 2020. To facilitate further interpretation, Figure 5.2

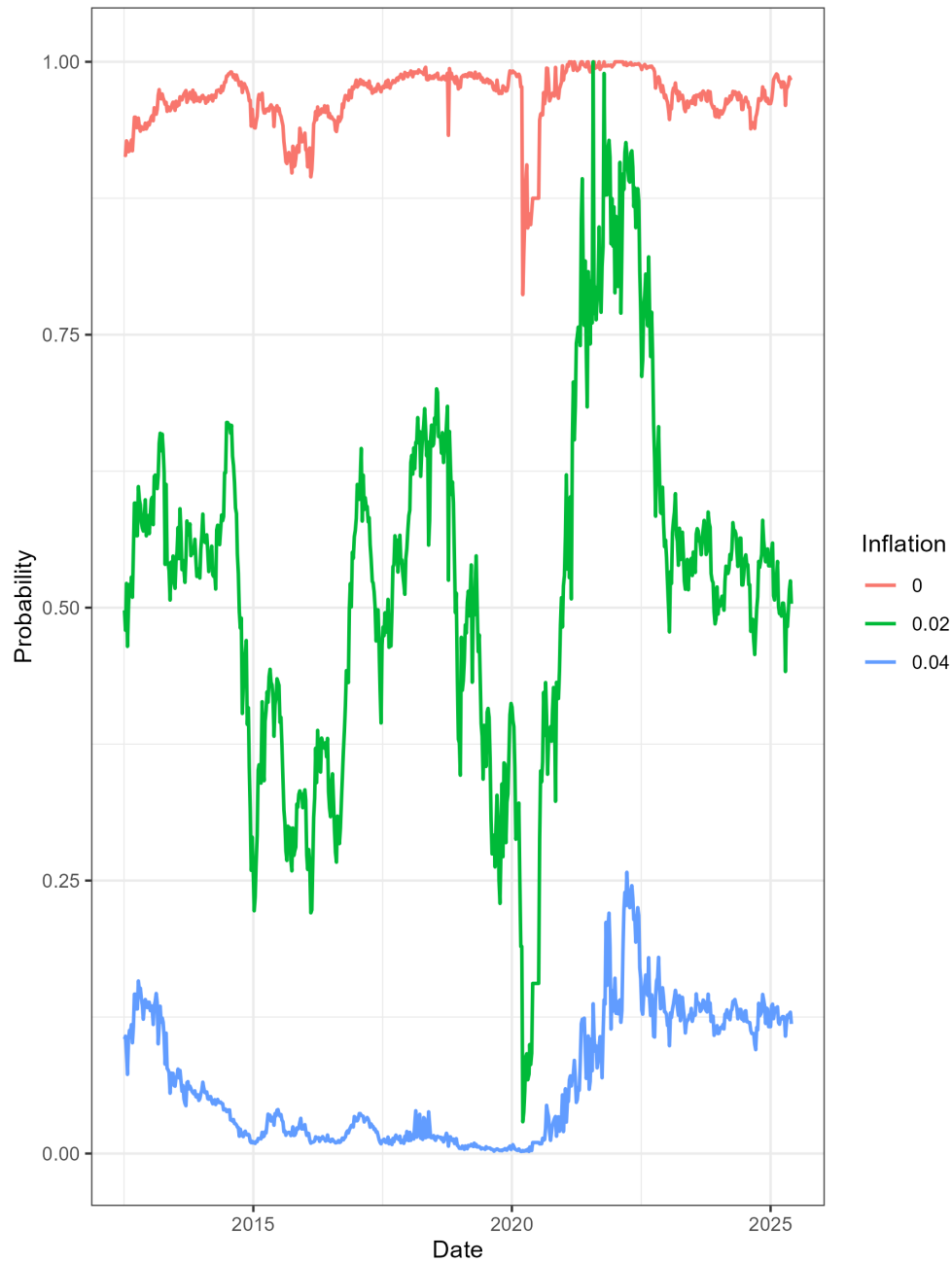
Figure 5.1: Fitted option-implied 5-year ahead risk neutral densities: weekly from July 5, 2012 through May 21, 2025



Source: Federal Reserve Bank of Minneapolis; Authors' calculations

extracts from these RNDs the probability of inflation being greater than or equal to selected levels of inflation over time. We now see more clearly that from 2020 the probability of inflation being greater than 4 percent increased and has remained elevated through 2025, even though the probability of inflation being greater than 2 percent fell back to its pre-pandemic values by October 2022.

Figure 5.2: Probability of five-years-ahead inflation being greater than selected levels based on the fitted option-implied risk neutral densities: weekly from July 5, 2012 through May 21, 2025



Source: Federal Reserve Bank of Minneapolis; authors' calculations

5.2.2 Macroeconomic surprises

We construct data on macroeconomic surprises computed as the difference between actual macroeconomic outcomes and expert forecasts, collected from Action Economics, LLC. This

results in the interpretation of a positive “surprise” as a case where the actual outcome was greater than expert forecaster expectations, and the opposite for a negative surprise. Our choice of macroeconomic surprises to include in our empirical analysis follows Kitsul and Wright (2013), which we then use a point of comparison to examine if the results of the functional analysis may differ than that of the simpler specification that they take.

These surprise data are mapped to the densities such that the surprise was observed in a short window prior to the release of the density in order to avoid endogeneity issues.⁶ In cases where no surprise data are mapped to a given density-week, the surprise data are set to zero.

Table 5.1 describes the nine macroeconomic surprise data series included in our analysis.⁷ Units of measurement for the macroeconomic surprises vary, which can complicate interpretation of our empirical results. To deal with this we standardize each series by demeaning and scaling by the standard deviation. Figure 5.3 plots the resulting time series of standardized macroeconomic surprises.

Large surprises relative to the series’ own history occurred for several series between 2020-2022, namely in nonfarm payrolls, retail sales, the fed funds rate, PPI, and CPI. Given the overlap with a period of changes in the location, scale, and skewness of the RNDs, we evaluate the extent to which these surprises have explanatory power associated with RND deviations from μ^* .

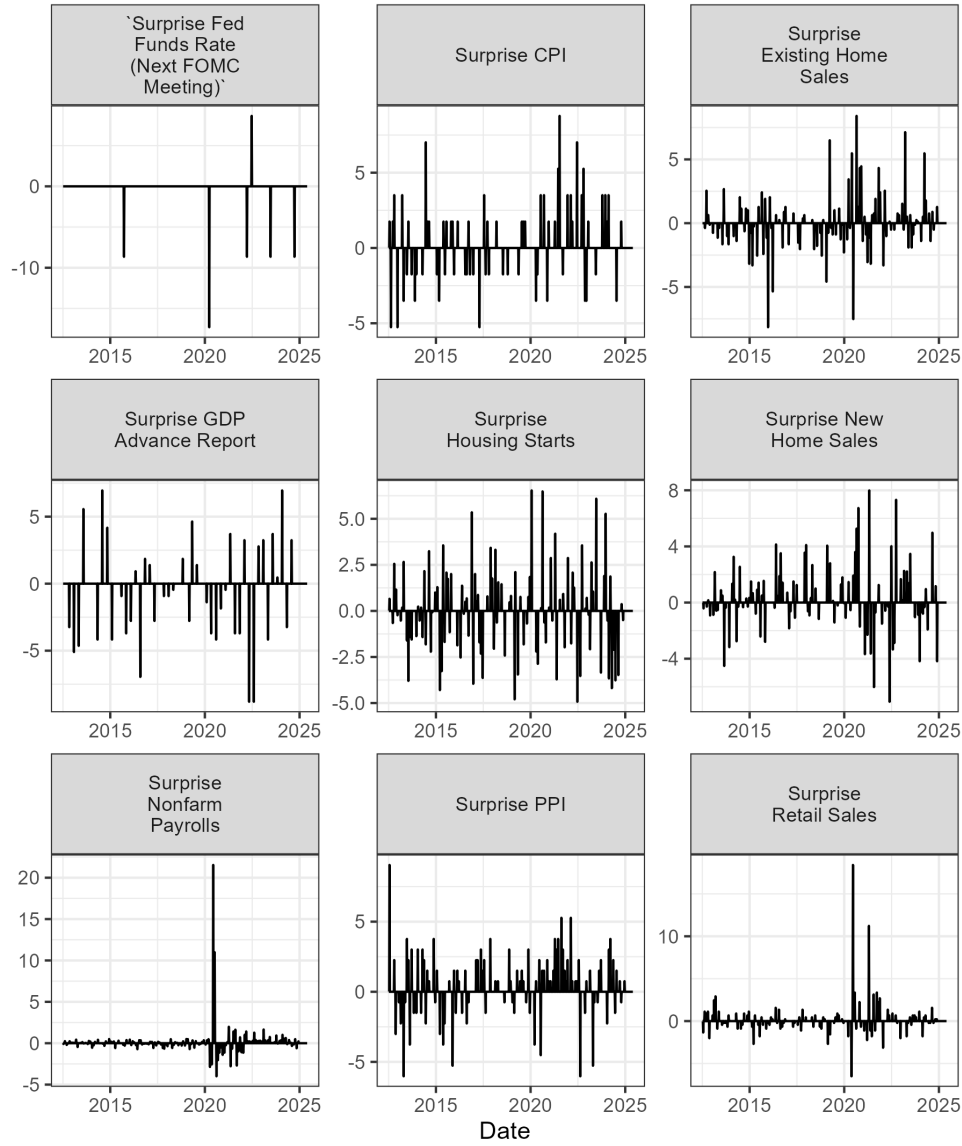
5.3 Probabilistic Anchoring of Inflation Expectations

Let $\pi_t^e(s)$ denote the probability density function characterizing market beliefs about long-run inflation, where s represents possible inflation outcomes. We observe $\pi_t^e(s)$ directly from our option-implied risk-neutral densities. We develop concepts to formalize what it means for $\pi_t^e(s)$ to be anchored, beginning with the policymaker’s “target density” for inflation.

⁶Specifically, we map a surprise to a given RND if it occurred at least 1 day prior to the publication date of the latest RND but not before the publication date of the previous RND.

⁷We provide summary statistics of the macroeconomic surprise data in Table 5.A.

Figure 5.3: Time series of macroeconomic surprises: weekly from from July 5, 2012 through May 21, 2025.



Source: Action Economics LLC, BEA, BLS, NAR, Federal Reserve Board of Governors, US Census, Authors' calculations. Each series is standardized to have mean 0 and standard deviation 1.

Definition 5.3.1 (Target Density). The policymaker has a target density $\mu^*(s)$ for long-run inflation, where:

$$\mu^*(s) : \mathcal{S} \rightarrow \mathbb{R}_+, \quad \int_{\mathcal{S}} \mu^*(s), ds = 1, \quad \int_{\mathcal{S}} s \cdot \mu^*(s), ds = \pi^*. \quad (5.1)$$

This density reflects not just the central bank's adopted target level $\pi^* = 2\%$, but the

Table 5.1: Description of macroeconomic surprises, observations from July 2012-May 2025. Source: Action Economics, Authors' calculations.

Indicator Name	Actual Source	Frequency	Units
CPI	BLS	Monthly	MoM percentage change
Existing Home Sales	NAR	Monthly	Millions of units
Fed Funds Rate (next FOMC meeting)	Federal Reserve	Eight per year	Percentage points (range midpoint)
GDP Advance Report	BEA	Quarterly	Annualized QoQ percentage change
Housing Starts	Census	Monthly	Millions of units
New Home Sales	Census	Monthly	Millions of units
Nonfarm Payrolls	BLS	Monthly	Thousands of USD
PPI	BLS	Monthly	MoM percentage change
Retail Sales	Census	Monthly	MoM percentage change

policymaker's tolerance for deviations around this value. While the Federal Reserve does not explicitly publish $\mu^*(s)$, we can infer it from policy communications and historical practice.

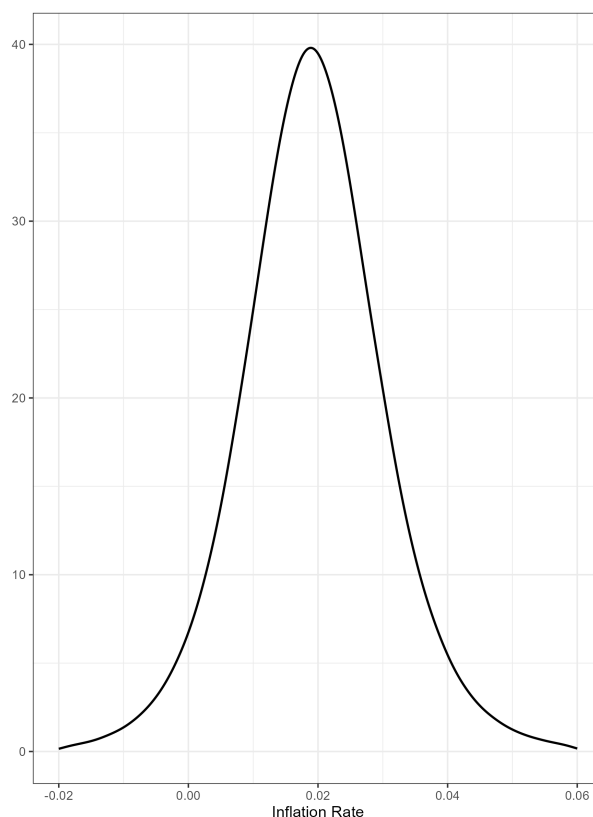
To estimate $\mu^*(s)$, we compute the average risk-neutral density over July 2012–July 2017, a period of broad consensus that expectations were well-anchored (Bundick and Smith, 2025). In the rolling regression framework that we use for analysis, this choice is also consistent with defining the initial condition with the target density. Because $\mu^*(s)$ is based on the risk-neutral densities, we assume that the policymaker's target density includes the risk-premium present in our risk-neutral densities. This is reasonable if we assume that policymakers care about both real-world beliefs and risk-premia present in financial markets when making policy decisions. This revealed-preference approach assumes that if expectations were anchored during this period, the average distribution reflects the market's understanding of the policymaker's acceptable tolerance for deviations.

Figure 5.4 shows the estimated $\mu^*(s)$: specifically, the inferred policymaker's target density has a mean of 2.0%, a standard deviation of 1.2%, skewness of 0.15, and kurtosis of 3.5.⁸ The fact that the density is centered at 2 percent confirms that agents did believe expectations were anchored around the FOMC's target. Then, the additional fact that the skewness and kurtosis estimates are close to the values implied by a Gaussian density, confirms that $\mu^*(s)$ is close to symmetric. This is consistent with the FOMC's 2017 "Statement on Longer-

⁸This implies that the policymaker is attributing a 66% probability to inflation between 1% and 3%, and a 92% probability to inflation between 0% and 4%.

Run Goals and Monetary Policy Strategy,” which emphasized “symmetric” concern about deviations of inflation from target.⁹

Figure 5.4: Policy mean function, derived from the empirical mean function over the period July 2012–July 2017.



Empirical estimates from structural models, such as DSGE models, suggest that during the Great Moderation period (1984–2004) – characterized by well-anchored expectations – US inflation also exhibited a standard deviation of approximately 1 percent on an annualized basis (Smets and Wouters 2007, Table 6). Reduced-form estimates (cf. the unobserved components model of Stock and Watson 2007) also put the standard deviation of inflation in this period at a similar value, reflecting transitory shocks with a SD of 0.6 and by 2004 a reduced SD of around 0.3 for the permanent component of inflation. These independent estimates of inflation uncertainty, over a period when inflation was deemed “anchored,” lend

⁹See https://www.federalreserve.gov/monetarypolicy/files/FOMC_LongerRunGoals_201701.pdf. It is also noteworthy that, even though realized inflation undershot the 2 percent target for much of 2012–2017, longer-run market expectations continued to average close to 2 percent.

additional support to our use of the density in Figure 5.4 as an estimate of the policymaker’s target density.

5.3.1 Two Dimensions of Probabilistic Anchoring

We extend traditional point-forecast anchoring tests to densities along two dimensions: whether expectations align with the target on average (unconditional), and whether they respond to incoming news (conditional).

Definition 5.3.2 (Unconditional Anchoring). Inflation expectations are *unconditionally anchored* if, on average, the expectations density equals the target density at each point on the support:

$$E[\pi_t^e(s)] = \mu^*(s), \quad \forall s \in \mathcal{S}. \quad (5.2)$$

This requires no *persistent* deviations from $\mu^*(s)$ at any inflation level. Temporary fluctuations around $\mu^*(s)$ are permitted, but systematic over- or under-estimation at specific inflation levels violates unconditional anchoring. Equation (5.2) generalizes the traditional test $E[\pi^e] = \pi^*$ (Kumar et al., 2015) from the mean to the entire distribution.

Definition 5.3.3 (Conditional Anchoring). Inflation expectations are *conditionally anchored* if, given macroeconomic information G_t , the density remains equal to the target:

$$E[\pi_t^e(s)|G_t] = \mu^*(s), \quad \forall s \in \mathcal{S}. \quad (5.3)$$

This generalizes Bernanke (2007)’s condition that mean expectations should be “insensitive to incoming data” to require orthogonality to news across the entire support. In news regressions (Gürkaynak et al., 2010; Kitsul and Wright, 2013; Nautz and Strohsal, 2015), G_t comprises macroeconomic surprises—the difference between actual data releases and prior forecasts. When combined with unconditional anchoring, conditional anchoring implies $\pi_t^e(s) = \mu^*(s) + \varepsilon_t(s)$ where $\varepsilon_t(s)$ is purely idiosyncratic.

Definition 5.3.4 (Perfect Probabilistic Anchoring). Inflation expectations exhibit *perfect probabilistic anchoring* when both unconditional anchoring (Definition 5.3.2) and conditional anchoring (Definition 5.3.3) hold. In this case, the entire distribution of beliefs matches the policymaker’s target and does not respond to macroeconomic news.

We refer to violations of perfect probabilistic anchoring as “unanchoring,” which our empirical framework decomposes into unconditional versus conditional sources.

5.3.2 Regional Anchoring: A Policy-Relevant Relaxation

Perfect anchoring as defined in Definition 5.3.4 is demanding—it requires expectations to match the target probabilistically at every inflation level. But, as discussed in the introduction, central banks often adopt tolerance ranges, recognizing that perfect control is neither feasible nor necessary. Adding to this, the FOMC’s September 2019 minutes discussed whether “a target range for inflation could be helpful in achieving the Committee’s objective of 2 percent inflation.”¹⁰ Research has also shown that asymmetric tolerance ranges can be optimal in the presence of constraints like the zero lower bound (Bianchi et al., 2021). We therefore introduce a weaker notion of probabilistic anchoring focusing only on policy-relevant regions.

Definition 5.3.5 (Tolerance Region). Define the tolerance region around the target as:

$$R_\varepsilon := [\pi^* - \varepsilon_L, \pi^* + \varepsilon_U] = [s_L, s_U]. \quad (5.4)$$

The choice of this region should reflect the specific (albeit perhaps assumed, if not stated publicly) preferences of the policymaker. In our empirical work, we consider a symmetric target, wherein $\varepsilon_L = \varepsilon_U = 2\%$, so $R_\varepsilon = [0\%, 4\%]$.

¹⁰See <https://www.federalreserve.gov/monetarypolicy/fomcminutes20190918.htm>.

Definition 5.3.6 (Regional Anchoring). Inflation expectations exhibit *regional anchoring* in R_ε if the probability mass in this region matches the target:

$$\begin{aligned} \text{Unconditional: } E \left[\int_{R_\varepsilon} \pi_t^e(s) ds \right] &= \int_{R_\varepsilon} \mu^*(s) ds =: \alpha^* \\ \text{Conditional: } E \left[\int_{R_\varepsilon} \pi_t^e(s) ds \middle| G_t \right] &= \alpha^*. \end{aligned} \tag{5.5}$$

Defining the coverage probability $\alpha_t := \int_{R_\varepsilon} \pi_t^e(s) ds$, regional anchoring requires $E[\alpha_t | G_t] = \alpha^*$ (conditional) or $E[\alpha_t] = \alpha^*$ (unconditional). Regional anchoring is less demanding than perfect anchoring: probability mass may shift *within* R_ε without violating regional anchoring, provided the total mass in R_ε matches α^* .

Interpreting violations of regional anchoring. When regional anchoring fails, not all deviations may be equally concerning from a policy perspective. If $E[\alpha_t] > \alpha^*$, the market concentrates *more* probability near the target than $\mu^*(s)$ prescribes—what we term “good regional anchoring.” This requires withdrawing probability from both tails, indicating lower disaster risk than the policymaker’s benchmark. While this violates perfect anchoring (Definition 5.3.4), it may be seen as benign from a policy perspective.

Conversely, if $E[\alpha_t] < \alpha^*$, excess probability mass resides in the tails. We distinguish:

- **Bad upper-tail anchoring:** $E \left[\int_{s_U}^\infty \pi_t^e(s) ds \right] > \int_{s_U}^\infty \mu^*(s) ds$ —heightened risk of persistent high inflation.
- **Bad lower-tail anchoring:** $E \left[\int_{-\infty}^{s_L} \pi_t^e(s) ds \right] > \int_{-\infty}^{s_L} \mu^*(s) ds$ —heightened deflation risk.

These concepts guide our interpretation of empirical results in Section 5.5: which regions of the distribution exhibit unanchoring, and whether violations represent heightened or diminished disaster risk relative to the policymaker’s tolerance.

5.4 Econometric Framework

To test probabilistic anchoring of inflation expectations as measured using the various definitions above, we use functional regression techniques discussed in the FDA literature. The functional component of this regression is the probability density function (PDF) capturing expected deviations of inflation from target:

$$\tilde{\pi}_t^e(s) = \pi_t^e(s) - \mu^*(s) = \beta_1(s) + \sum_{j=2}^p x_{tj} \beta_j(s) + \varepsilon_t(s) \quad \forall s \in \mathcal{S} \quad (5.6)$$

where x_{tj} are standardized macroeconomic surprises (the difference between actual releases and forecasts), and $\beta_j(s)$ are functional coefficients. The functional intercept $\beta_1(s)$ captures persistent deviations from $\mu^*(s)$ (unconditional unanchoring), while the slope coefficients $\beta_j(s)$ capture the response to macro news (conditional unanchoring).

Our use of macroeconomic surprises rather than levels to model the expected inflation density relative to its steady-state value addresses potential simultaneity bias. Structural models (Carvalho et al., 2023; Gáti, 2023) feature two-way causality between current inflation and long-run expectations: short-term surprises affect long-run beliefs, which then feed back into current inflation via the expectations-augmented Phillips curve.

While current inflation may depend on long-run expectations via the expectations-augmented Phillips curve and long-run expectations can exhibit endogenous sensitivity to short-term inflation surprises, surprises are by definition orthogonal to agents' expectations formed prior to the data release. Moreover, by specifying our regression as deviations from the target density $\mu^*(s)$, we effectively control for the steady-state level, similar to how first-differencing removes a unit root, while preserving information about persistent deviations captured by $\beta_1(s)$.

PDFs are subject to non-negativity and unit integral constraints, all of which must be carefully treated when using traditional FDA methods to retain the properties of the PDF. Relative to our analysis, linear operations on density objects require transformation to a

Hilbert space in order for the result to be a valid density. Petersen and Müller (2016) suggest a general framework for FDA on PDFs by choosing a suitable transformation method that allows for linear operations and back-transformation to the original density space.

Recent papers in the economic literature have used the log quantile density (LQD) transformation (Clements and Wang (2024); Huber et al. (2024b)). We use an alternative transformation based on results from Hron et al. (2016) and Talská et al. (2018), who perform linear operations in the Bayes Space using the centered log ratio (CLR) transformation. Given our interests in this analysis, we choose the CLR transformation because it allows for straightforward back-transformation of the densities we work with *and* for the coefficients that we estimate in our functional regressions, as shown in Talská et al. (2018).

Our econometric methods utilizes multiple linear operations on density objects. First, we subtract the policy maker’s time-invariant prior $\mu^*(s)$ from the densities so that we explain deviations from this prior. Next, we apply the Karhunen-Loève expansion to reduce the infinite-dimensional densities to a finite number of functional scores and eigenfunctions that explain the majority of the total variation in the deviations. We then regress the functional scores on a vector of scalar macroeconomic surprises using a Bayesian linear regression. Finally we construct functional coefficients and predicted densities from the estimated coefficients and predicted scores. The following section describes how we achieve these aims.

5.4.1 Bayes Space and Centered Log Ratio Transformation

We utilize the geometry of the Bayes space $\mathcal{B}^2$ to perform linear operations on PDFs. This approach is motivated by the idea that densities contain relative information and thus represent compositional data, that is, the unit integral constraint suggests that an increase in probability mass in one are of the density requires a decrease in probability mass somewhere else. As a result, the absolute difference between two densities will generally not result in a valid density. A more appropriate transformation is multiplicative, so that probability mass

is redistributed across the support via pointwise ratios.

The Bayes space $\mathcal{B}^2$ provides a geometry for multiplicative operations on compositional data under an appropriate transformation of the data, and probability density functions can be considered as infinite-dimensional compositional data (Egozcue et al. (2006); Van den Boogaart et al. (2014)). Multiplicative operations on densities in the Bayes space result in valid densities, and include *perturbation* and *powering*. These operations motivate the name of the Bayes space because they mimic Bayesian updating. For f and g in $\mathcal{B}^2$, $s \in I = [a, b]$, perturbation is the analogue of addition in the Bayes space, defined as

$$f \oplus g = \frac{f(s) g(s)}{\int f(u) g(u) du} \quad (5.7)$$

For $c \in \mathbf{R}$, powering is the Bayes space analogue of multiplication, defined as

$$f \odot c = \frac{f(s)^c}{\int f(u)^c du} \quad (5.8)$$

The inner product in $\mathcal{B}^2$ is

$$\langle f, g \rangle_{\mathcal{B}^2} = \frac{1}{2(b-a)} \int_a^b \int_a^b \ln \frac{f(s)}{f(u)} \cdot \ln \frac{g(s)}{g(u)} ds du \quad (5.9)$$

and the norm and distance measures are

$$\|f\|_{\mathcal{B}^2} = \sqrt{\langle f, f \rangle_{\mathcal{B}^2}} \quad (5.10)$$

$$d_{\mathcal{B}^2}(f, g) = \|f \ominus g\|_{\mathcal{B}^2} \quad (5.11)$$

with $(f \ominus g)(t) = [f \oplus (-1) \odot g](t)$ representing perturbation-subtraction.

In both cases, the renormalization term in each expression returns a valid density so that probability mass is redistributed across the support as a result of the operations. Importantly, the identity element e of perturbation that satisfies $f \oplus e = f$ is the uniform density $e(s) =$

$\frac{1}{b-a}$.¹¹. This holds because

$$f \oplus e = \frac{ef(s)}{\int ef(u) du} = \frac{f(s)}{\int f(u) du} = f(s) \quad (5.12)$$

since $f(u)$ already integrates to 1.

To perform operations in $\mathcal{B}^2$, we transform the densities using the centered log ratio (CLR) transform (Van den Boogaart et al. (2014)). The CLR transformation is an isometric isomorphism between $\mathcal{B}^2$ and the space L^2 of square-integrable real measurable functions, meaning inner products and distances are preserved when transforming between spaces. This transformation allows us to perform linear operations such as functional principal component analysis (FPCA) and linear regression. In doing so, we combine the FPCA and functional regression analysis methods of Huber et al. (2024a) with the CLR-based functional regression methods of Talská et al. (2018).

Denoting $clr[\cdot]$ as a functional object transformed to the CLR space, the CLR is defined for a functional composition f as

$$clr[f](s) = \ln f(s) - \frac{1}{b-a} \int \ln f(u) du \quad (5.13)$$

Densities that result from the CLR transformation have a zero integral constraint on the support I , so that $\int_I clr[f](s)ds = 0$. A CLR density object can be inverted back to the density space,

$$f(s) = clr^{-1}[f](s) = \exp(clr[f](s)) \quad (5.14)$$

The CLR transformation linearizes the perturbation and powering operations in the Bayes Space so that they can be represented as addition and multiplication. Perturbation under a

¹¹In L^2 space, the analogue of this statement would state that the identity element of addition is $e = 0$.

CLR transformation becomes

$$clr[f \oplus g] = clr[f(s)] + clr[g(s)] \quad (5.15)$$

Powering under a CLR transformation becomes

$$clr[f \odot g] = \alpha \cdot clr[f(s)] \quad (5.16)$$

5.4.2 Testing Probabilistic Anchoring via Functional Regression

Having established the CLR transformation framework, we now develop our empirical strategy for testing probabilistic anchoring. Our approach centers on decomposing deviations from the target density into systematic versus idiosyncratic components, with the systematic component further split into unconditional and conditional sources.

The Core Decomposition

Following Talská et al. (2018), we write the function-on-scalar regression (FoSR) in Bayes space using perturbation ($\oplus$) and powering ($\odot$) as:

$$\tilde{\pi}_t^e(s) = \bigoplus_{j=1}^p (x_{tj} \odot \beta_j)(s) \oplus \varepsilon_t(s) \quad (5.17)$$

where $\bigoplus$ is the $\mathcal{B}^2$ analogue of $\sum$ in L^2 , $\tilde{\pi}_t^e(s) = \pi_t^e(s) \ominus \mu^*(s)$ represents the perturbation-difference between the observed density and the target density, x_{tj} are macroeconomic surprises (with $x_{t1} = 1$ for the intercept), $\beta_j(s)$ are functional coefficients in density space, and $\varepsilon_t(s)$ is the error density.

Applying the CLR transformation to this equation linearizes the Bayes space operations, yielding our core decomposition in L^2 space:

$$\text{clr}[\tilde{\pi}_t](s) = \underbrace{\text{clr}[\beta_1](s)}_{\text{Unconditional}} + \underbrace{\sum_{j=2}^p x_{tj} \text{clr}[\beta_j](s)}_{\text{Conditional}} + \underbrace{\text{clr}[\varepsilon_t](s)}_{\text{Idiosyncratic}} \quad (5.18)$$

This decomposition maps directly to our anchoring concepts from Section 5.3.1:

- $\text{clr}[\beta_1](s)$ captures unconditional unanchoring (Definition 5.3.2): time-invariant deviations from $\mu^*(s)$ at inflation level s . When $\text{clr}[\beta_1](s) \neq 0$, expectations persistently deviate from target.¹²
- $\sum_{j=2}^p x_{tj} \text{clr}[\beta_j](s)$ captures conditional unanchoring (Definition 5.3.3): sensitivity to macroeconomic surprises at inflation level s . When $\text{clr}[\beta_j](s) \neq 0$ for $j \geq 2$, news affects the distribution.
- $\text{clr}[\varepsilon_t](s)$ represents purely idiosyncratic variation consistent with perfect probabilistic anchoring (Definition 5.3.4).

Quantifying Unanchoring: Functional R^2 Decomposition

We quantify the importance of unanchoring using functional R^2 measures that partition explained variation at each inflation level s :

$$R^2(s) = \frac{\text{SSY}(s) - \text{SSE}(s)}{\text{SSY}(s)} \quad (5.19)$$

where the functional sum of squared errors (SSE) and total sum of squares (SSY) are defined as $\text{SSY}(s) = \sum_t \text{clr}[\tilde{\pi}_t](s)^2$ and $\text{SSE}(s) = \sum_t (\text{clr}[\tilde{\pi}_t](s) - \widehat{\text{clr}}[\tilde{\pi}_t](s))^2$. Under perfect probabilistic anchoring (Definition 5.3.4), $R^2(s) = 0$ everywhere—the target density $\mu^*(s)$ fully explains observed expectations.

We decompose $R^2(s)$ into components corresponding to equation (5.18):

¹²Because $\mu^*(s)$ is estimated from a fixed period (2012–2017), the intercept β_1 reflects deviations from this benchmark. Alternative specifications of $\mu^*(s)$ would shift β_1 but not affect the slope coefficients $\beta_2, \dots, \beta_p$. We assess robustness to alternative target densities in Section 5.5.4.

Unconditional $R^2(s)$: Variation explained by persistent deviations (β_1 alone), computed by setting $\beta_j = 0$ for $j \geq 2$. This quantifies violations of unconditional anchoring (Definition 5.3.2) at inflation level s : when $R_{\text{uncond}}^2(s) > 0$, expectations systematically deviate from $\mu^*(s)$ independent of incoming news.

Conditional $R^2(s)$: Variation explained by macro surprises ($\beta_2, \dots, \beta_p$), computed by setting $\beta_1 = 0$. This quantifies violations of conditional anchoring (Definition 5.3.3) at inflation level s : when $R_{\text{cond}}^2(s) > 0$, expectations respond to news at that inflation level.

Partial $R_j^2(s)$: Variation uniquely explained by surprise j , computed via standard partial regression formulas. This identifies which macroeconomic indicators drive conditional unanchoring at each inflation level.

Conversely, $R^2(s) > 0$ indicates deviations from anchoring, with the decomposition into unconditional and conditional components revealing *why* anchoring fails: persistent shifts in beliefs (β_1) versus sensitivity to incoming data ($\beta_2, \dots, \beta_p$).

The **global R^2** summarizes total unanchoring across the entire distribution, averaging across the support and weighting by each inflation level's share of total variation:

$$R^2 = \frac{\int_{\mathcal{S}} [\text{SSY}(s) - \text{SSE}(s)] ds}{\int_{\mathcal{S}} \text{SSY}(s) ds} = \int_{\mathcal{S}} R^2(s) \cdot w(s) ds \quad (5.20)$$

where $w(s) = \text{SSY}(s) / \int_{\mathcal{S}} \text{SSY}(s) ds$ weights each inflation level by its share of total variation. Regions of the distribution with more variation in deviations from $\mu^*(s)$ receive proportionally more weight in the global measure. This measures violations of perfect probabilistic anchoring (Definition 5.3.4) pooling both unconditional and conditional sources—what we term “total unanchoring.”¹³

¹³We can also compute functional F-statistics to test if the intercept β_1 and macroeconomic surprise coefficients $\beta_2, \dots, \beta_p$ jointly explain variation across the entire distribution. Under the null hypothesis of perfect probabilistic anchoring (Definition 5.3.4), the test statistic is $F = \text{MSR}/\text{MSE} \sim F(p, N - p)$, where $\text{MSR} = \int_{\mathcal{S}} [\text{SSY}(s) - \text{SSE}(s)] ds / df_{\text{model}}$ is the mean square for regression and $\text{MSE} = \int_{\mathcal{S}} \text{SSE}(s) ds / df_{\text{error}}$ is the mean square for error, with $df_{\text{model}} = p$ and $df_{\text{error}} = N - p$. Pointwise F-statistics $F(s) = \text{MSR}(s)/\text{MSE}(s)$ can also be computed to explore local effects at each inflation level s , though these are subject to multiple testing

The functional R^2 framework has two key advantages for our purposes. First, it directly quantifies the *economic importance* of unanchoring—what fraction of distributional variation is systematic rather than noise? Second, the decomposition into unconditional versus conditional components reveals why anchoring fails at each point on the distribution.

Locating Unanchoring: Functional Coefficients

While $R^2(s)$ quantifies how much unanchoring occurs at each inflation level, the functional coefficients $\beta_j(s)$ reveal how macroeconomic surprises redistribute probability mass. We interpret coefficients in density space by back-transforming via:

$$\beta_j(s) = \text{clr}^{-1}[\beta_j](s) = \frac{\exp(\text{clr}[\beta_j](s))}{\int_I \exp(\text{clr}[\beta](u)) du} \quad (5.21)$$

The back-transformed coefficient $\beta_j(s)$ represents a multiplicative reweighting kernel showing how probability mass is redistributed in response to surprise j :

- When $\text{clr}[\beta_j](s) = 0 \Rightarrow \beta_j(s) = e(s) = \frac{1}{b-a}$ (the neutral density): no reweighting of probability mass at s .
- When $\text{clr}[\beta_j](s) > 0 \Rightarrow \beta_j(s) > e(s)$: probability mass is pulled *toward* inflation level s .
- When $\text{clr}[\beta_j](s) < 0 \Rightarrow \beta_j(s) < e(s)$: probability mass is pushed *away from* inflation level s .

Because densities integrate to one, $\beta_j(s)$ describes a zero-sum redistribution: regions where $\beta_j(s) > e(s)$ gain probability mass at the expense of regions where $\beta_j(s) < e(s)$.

The functional coefficients thus reveal not just *whether* macro surprises affect expectations (as in point regressions), but *where* on the distribution they matter—center versus tails versus specific inflation levels.

concerns and should be interpreted as exploratory.

Regional Tests

The functional R^2 measures and coefficient plots assess anchoring pointwise at each inflation level s . However, as discussed in Section 5.3.2, central banks may tolerate deviations from the target at specific inflation levels provided the total probability mass within a tolerance region matches the target. We therefore test regional anchoring (Definition 5.3.6) by computing credible intervals for the integrated coefficients within $R_\epsilon = [0\%, 4\%]$:

$$\text{Unconditional: } \int_{R_\epsilon} \text{clr}[\beta_1](s) ds \quad (5.22)$$

$$\text{Conditional: } \int_{R_\epsilon} \text{clr}[\beta_j](s) ds, \quad j = 2, \dots, p \quad (5.23)$$

The unconditional test (5.22) asks whether there are persistent shifts in probability mass within R_ϵ relative to the target $\mu^*(s)$ —that is, violations of unconditional anchoring (Definition 5.3.2) aggregated over this region. The conditional tests (5.23) ask whether macroeconomic surprises systematically shift probability mass into or out of R_ϵ —that is, violations of conditional anchoring (Definition 5.3.3) aggregated over this region. When credible intervals exclude zero, we reject regional anchoring in that dimension.

These regional tests are less demanding than pointwise tests: probability mass may shift within R_ϵ (e.g., from 1.5% to 2.5%) without violating regional anchoring, provided the total mass in the region matches the target. Combined with analogous tests on the tail regions ($s < 0\%$ and $s > 4\%$), they reveal whether violations represent good regional anchoring (excess mass concentrated near the target) or bad tail anchoring (excess mass in disaster regions).

5.4.3 Functional Principal Components Analysis

The appeal of FoSR for our study is the ability to regress an infinite-dimensional functional object on a vector of covariates to examine marginal effects and the proportion of varia-

tion explained across the support of the function. Functional principal component analysis (FPCA) offers a computationally efficient way to reduce the infinite-dimensional densities to a finite number of basis functions and scores that explain most of the variation.

The results of Hron et al. (2016) show that performing FPCA on CLR-transformed densities allows for use of standard FPCA methods. We apply the Karhunen-Loève expansion, but substitute the policymaker's target density $\mu^*(s)$ for the empirical mean typically used in FPCA methods, so that

$$\text{clr}[\pi_t^e](s) = \text{clr}[\mu^*](s) + \sum_{k=1}^{\infty} \xi_{kt} \phi_k(s) \quad (5.24)$$

where ξ_{kt} are the functional principal component scores and $\phi_k(s)$ are the eigenfunctions. We choose the number of components K to explain 95 percent of the total variation in the CLR-transformed densities.¹⁴ This gives a finite approximation:

$$\text{clr}[\pi_t^e](s) = \text{clr}[\mu^*](s) + \sum_{k=1}^K \xi_{kt} \phi_k(s) + e_t(s) \quad (5.25)$$

Given this finite representation, FPCA offers a parsimonious way to estimate the FoSR. To model the marginal impacts of the macroeconomic surprises on the densities $\pi_t^e(s)$, we regress each of the K scores ξ on the design matrix X , collapsing the estimation to a simple linear regression:

$$\xi_{kt} = \beta_{1k} + \sum_{j=2}^p \text{clr}[\beta_{jk}] x_{jt} + \varepsilon_{kt} \quad (5.26)$$

The eigenfunctions $\phi_k(s)$ allow us to expand the estimations across the functional support. To construct the functional coefficient for covariate j , we aggregate across components:

$$\text{clr}[\hat{\beta}_j](s) = \sum_{k=1}^K \widehat{\text{clr}[\beta_{jk}]} \phi_k(s) \quad (5.27)$$

¹⁴We also examined using 99 percent as the threshold. This did not materially affect our results, as the majority of variation is typically captured by the first three components.

Similarly, the predicted scores yield predicted densities:

$$\hat{\xi}_{kt} = \sum_{j=1}^p \widehat{\text{clr}[\beta_{jk}]} x_{jt} \quad (5.28)$$

$$\widehat{\text{clr}[\tilde{\pi}_t^e]}(s) = \sum_{k=1}^K \hat{\xi}_{kt} \phi_k(s) \quad (5.29)$$

This ties our empirical model back to the decomposition in equation (5.18).

Bayesian Estimation

We estimate (5.26) using standard Bayesian normal linear regression with independent normal-gamma priors. We sample from the conditional posteriors using the following Gibbs sampling algorithm:

1. Set initial values for β and $h = 1/\sigma^2$.
2. Draw $\beta|\xi, h \sim N(\bar{\beta}, \bar{V})$.
3. Draw $h|\xi, \beta \sim G(\bar{s}^{-2}, \bar{\nu})$.
4. Repeat steps 2–3.

The posterior parameters are:

$$\bar{V} = (\underline{V}^{-1} + hX'X)^{-1}, \quad (5.30)$$

$$\bar{\beta} = \bar{V}(\underline{V}^{-1}\underline{\beta} + hX'\xi), \quad (5.31)$$

$$\bar{\nu} = N + \underline{\nu}, \quad (5.32)$$

$$\bar{s}^2 = \frac{(\xi - X\beta)'(\xi - X\beta) + \underline{\nu}s^2}{\bar{\nu}}. \quad (5.33)$$

We assume a flat prior on $p(\beta)$ and the Jeffreys prior $p(h) \propto h^{-1}$. For numerical implementation, the flat prior on β is approximated using $\underline{\beta} = 0$ and $\underline{V} = 100I$. The Jeffreys prior on $p(h)$ takes $\underline{\nu} \rightarrow 0$, so that $h|\xi, \beta \sim G(N, (\xi - X\beta)'(\xi - X\beta))$.

For point estimates of $\beta_j(s)$, we apply equation (5.21) to the draws retained after burn-in and compute the posterior means of the back-transformed estimates. Inference on $\text{clr}[\beta_j](s)$ requires adjustment for serial correlation in the residuals, which would otherwise produce credible intervals that are too narrow. Following Müller (2013), we adjust the posterior covariance matrix using the HAC estimator $\hat{V}_{\text{HAC}}$ of Newey and West (1987), so that $\text{clr}[\beta] \sim N(\hat{\beta}, \hat{V}_{\text{HAC}})$. To construct credible intervals for $\beta_j(s)$, we replace the posterior covariance with the sandwich covariance and then apply equation (5.21) to the adjusted draws.¹⁵

Our Gibbs sampling algorithm takes 50,000 draws and discards the first 10,000 as burn-in to construct posterior mean point estimates and 95 percent credible intervals. Trace plots for each regression on score k (not reported) confirm good convergence of the sampler.

Implementation Details

For computational purposes, we discretize the support $\mathcal{S} = [-0.02, 0.06]$ at $N = 801$ evenly-spaced grid points $\{s_1, s_2, \dots, s_N\}$ in increments of 1 basis point. This yields inflation levels ranging from -2% to 6% , matching the range of available strike prices in the inflation options data. The neutral density for this support is $e(s) = 1/(0.06 - (-0.02)) = 12.5$, which serves as the benchmark when interpreting back-transformed coefficients $\beta_j(s)$ (see (5.21)).

In matrix form, the CLR-transformed densities $\text{clr}[\tilde{\pi}_t^e]$ form a $T \times N$ matrix, where T is the number of time periods in the estimation window and each row represents a density evaluated at the N grid points. The coefficient functions $\text{clr}[\beta](\cdot)$ form a $p \times N$ matrix, where p is the number of regression parameters (one intercept plus $p - 1$ macroeconomic surprise variables). The error function $\text{clr}[\varepsilon](\cdot)$ is similarly a $T \times N$ matrix with $\text{clr}[\varepsilon](\cdot) \sim (0, \Sigma)$.

We estimate equation (5.26) separately for each of the K functional principal component scores using the Bayesian linear regression described above. For each score k , we obtain posterior distributions for the coefficients $\{\text{clr}[\beta_{jk}]\}_{j=1}^p$, which we aggregate across components

¹⁵We construct $\hat{V}_{\text{HAC}}$ using standard methods. The long-run variance of the moment contributions $g_t = x_t \hat{u}_t$ is estimated as $\hat{S} = \hat{\Gamma}_0 + \sum_{\ell=1}^L w_\ell (\hat{\Gamma}_\ell + \hat{\Gamma}_\ell')$, where $\hat{\Gamma}_\ell = \frac{1}{T} \sum_{t=\ell+1}^T g_t g_{t-\ell}'$. We use a Bartlett kernel for w_ℓ and automatic bandwidth selection for L . Then $\hat{Q} = \frac{1}{T} X'X$, and $\hat{V}_{\text{HAC}} = \frac{1}{T} \hat{Q}^{-1} \hat{S} \hat{Q}^{-1}$.

via equation (5.27) to obtain the functional coefficients $\text{clr}[\beta_j](s)$. Back-transformation to density space via equation (5.21) yields the interpretable coefficients $\beta_j(s)$. Credible intervals for hypothesis tests and regional tests are constructed using the HAC-adjusted posterior covariance matrix $\hat{V}_{\text{HAC}}$.

5.5 Empirical Results: Three Findings About Probabilistic Anchoring

We estimate the functional regression (5.6) over rolling five-year windows, advancing one month at a time. Our first estimation window spans July 2012–July 2017; we then roll forward maintaining a five-year window length through May 2025. In each window, we select the number of functional principal components K to explain 95% of the variation in CLR-transformed densities (see).¹⁶ This rolling approach allows us to trace how the degree and nature of probabilistic anchoring evolved over time, rather than imposing a single estimate over the full sample period.

Our analysis reveals three key findings that traditional point-forecast metrics would miss.

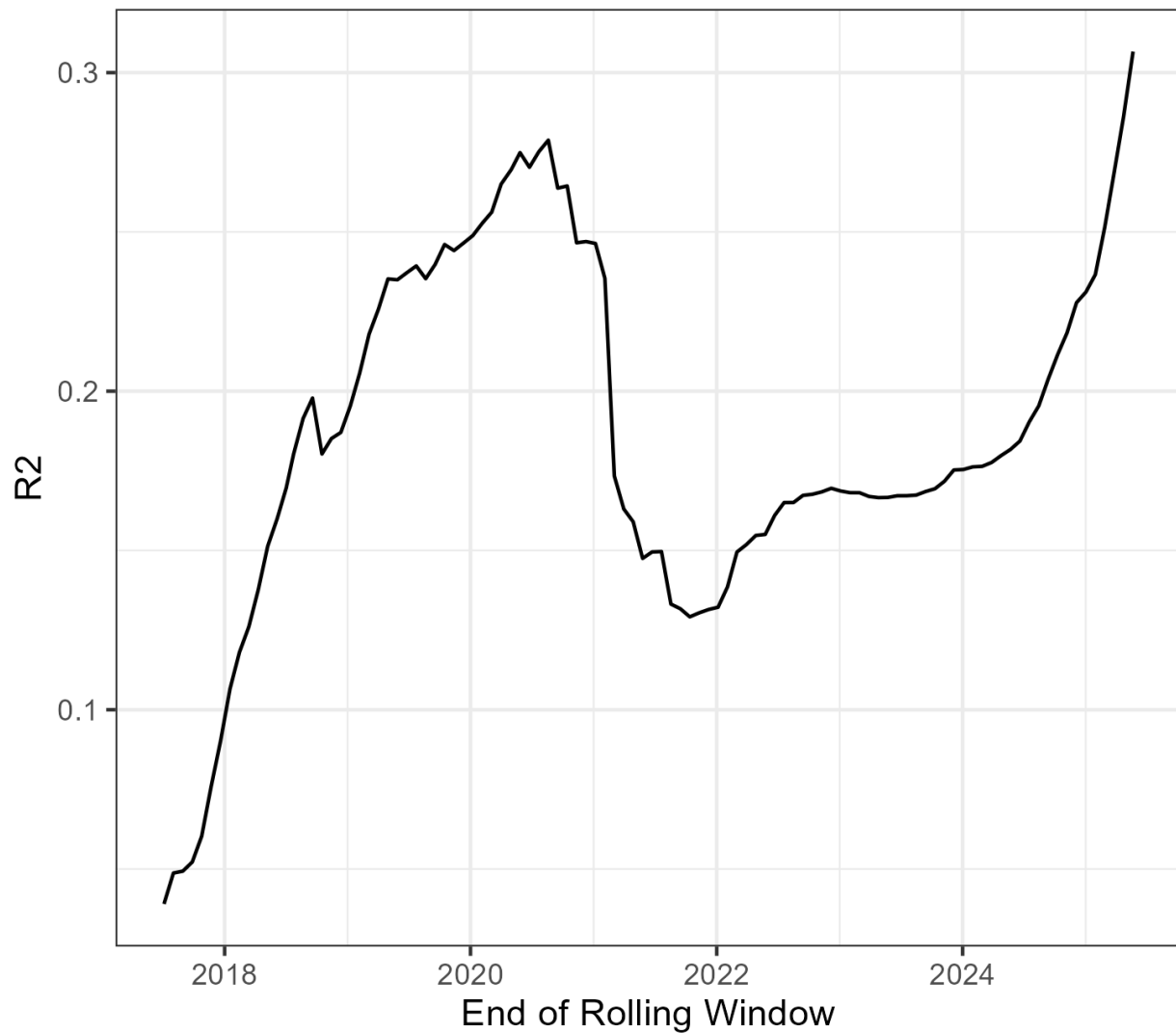
5.5.1 Fact 1: Total Anchoring Deteriorated in Two Waves

Figure 5.5 plots the global R^2 from equation (5.20), which measures *total unanchoring*—violations of perfect probabilistic anchoring (Definition 5.3.4) pooling both unconditional and conditional sources. Three distinct periods emerge:

2017–2018: Strong anchoring. The global R^2 , (5.20), remains near zero through late-2017, indicating the entire distribution of expectations matches the target density $\mu^*(s)$. Deviations from $\mu^*(s)$ are largely idiosyncratic—neither the residual intercept β_1 nor macroe-

¹⁶This typically yields $K = 3$ – 6 components. Using 99% as the threshold does not materially affect our results, as most variation is captured by the first three components.

Figure 5.5: Rolling estimates of total unanchoring: the global R^2 , (5.20)



Notes: Underlying functional regressions are estimated at monthly frequency over rolling 5-year windows. The x-axis represents the end-date of each rolling estimation window.

conomic surprises explain systematic variation. In the language of Definition 5.3.4, expectations exhibit perfect probabilistic anchoring during this period.

2018–2020: First wave of unanchoring. The global R^2 rises sharply from near zero in late-2017 to peak around 0.27 in mid-2020. During this period, over one-quarter of the variation in deviations from $\mu^*(s)$ can be explained by our model, indicating systematic (rather than idiosyncratic) departures from the target density. We reject the null hypothesis

of perfect probabilistic anchoring from 2018 onward (see Figure 5.B in the online appendix for the global functional F-test results).

2021–2025: Brief re-anchoring then second wave. Anchoring improves through 2021–2023 as the global R^2 declines, before unanchoring resurges from 2024. By the final window ending May 2025, the global R^2 exceeds 0.30—the highest level in our sample—indicating that nearly one-third of distributional deviations from target are now systematic.

While Figure 5.5 documents *when* anchoring deteriorated, it does not reveal *why*—whether due to persistent shifts in beliefs (unconditional unanchoring) or sensitivity to news (conditional unanchoring)—nor *where* on the distribution these failures occur. We address these questions in the next two subsections.

5.5.2 Fact 2: The Nature of Unanchoring Changed from Good to Bad

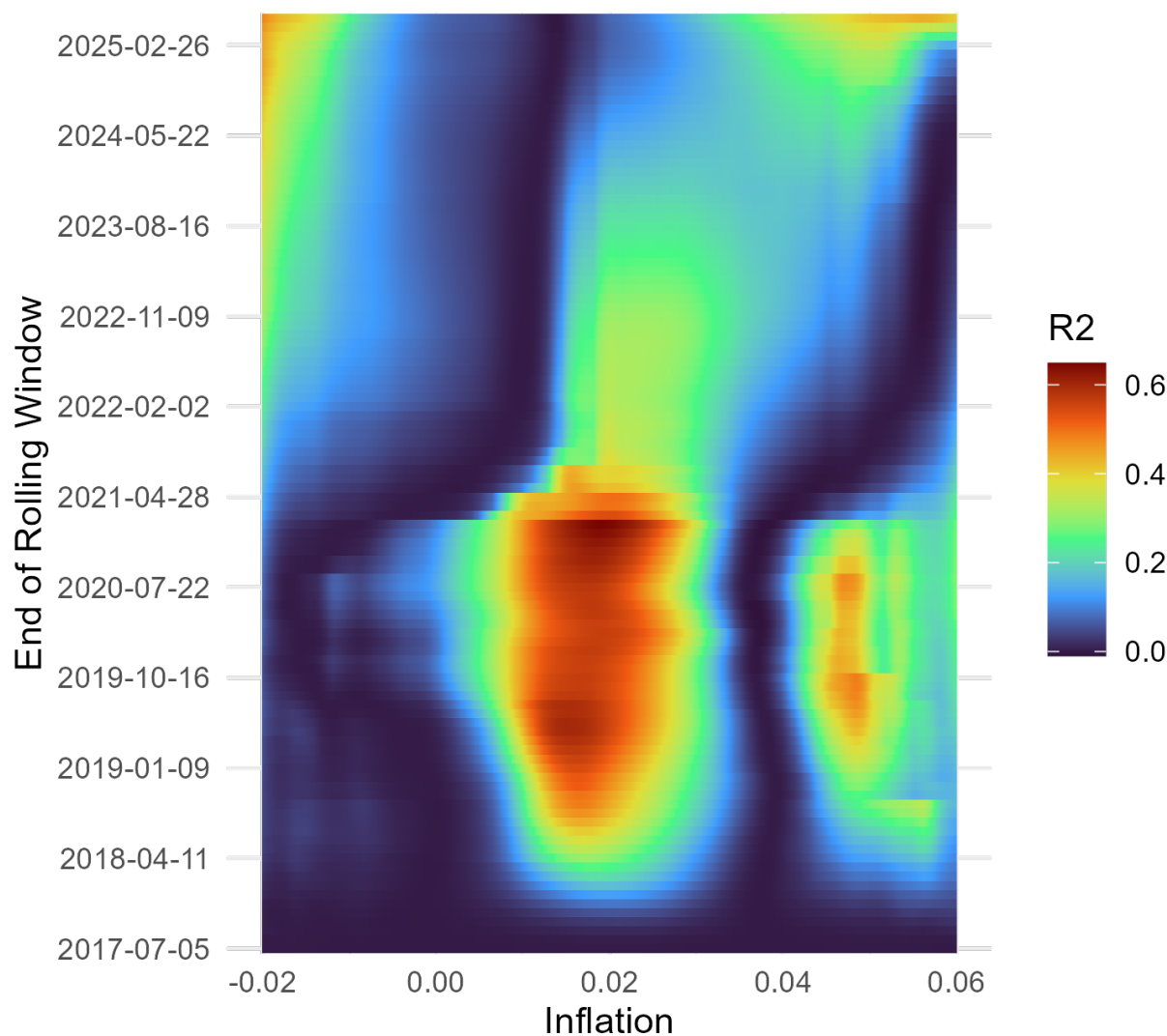
The two waves of unanchoring in Figure 5.5 reflect fundamentally different phenomena. We demonstrate this by decomposing total unanchoring into: (i) which regions of the distribution exhibit systematic deviations from anchoring (Figure 5.6), (ii) whether this reflects unconditional shifts versus news sensitivity (Figure 5.7), and (iii) whether violations represent good or bad anchoring (Figure 5.8).

Where unanchoring occurs: Center versus tails

Figure 5.6 plots the functional $R^2(s)$ from equation (5.19) across the support of the inflation expectations density. The figure reveals a striking shift in *where* deviations from $\mu^*(s)$ are concentrated:

2018–2020: Center unanchoring. The evidence of unanchoring around 2020 occurred primarily in the central region of the densities. Systematic deviations from $\mu^*(s)$ are concen-

Figure 5.6: Functional $R^2(s)$ from rolling regressions. Darker shading indicates higher explained variation at inflation level s .



trated when inflation expectations are between 1 and 3 percent, with $R^2(s)$ values exceeding 0.4 near the 2% target. This means our model—the intercept β_1 and macro surprises—explains over 40% of the variation in how the expectations density deviates from the policymaker’s target at these inflation levels. In contrast, the tails (below 0% and above 4%) exhibit low R^2 values, indicating these regions remained well-anchored.¹⁷

¹⁷We can again supplement inspection of the functional R^2 s with functional F-tests. Figure 5.C shows the global p-values associated with each of the pointwise F-tests over the rolling results, respectively. These F-tests are based on the full model, which includes both tests of conditional and unconditional unanchoring. For the pointwise results, the p-values are relatively high initially but declining sharply across most of the

2024–2025: Tail unanchoring. While the center region generally dominates the high R^2 values prior to 2021, there is a rightward migration of relatively higher R^2 values beginning in 2021. From 2024, the rising global R^2 values are instead driven by the upper and lower tail regions, with very little explanatory power in the central region. The $R^2(s)$ surface shows high values for $s < 0\%$ and $s > 4\%$, while the center near 2% exhibits $R^2(s) \approx 0$.

Unconditional versus conditional sources

Figure 5.7 decomposes the functional R^2 into the component explained by macroeconomic surprises alone (excluding the intercept β_1), revealing the role of conditional unanchoring.

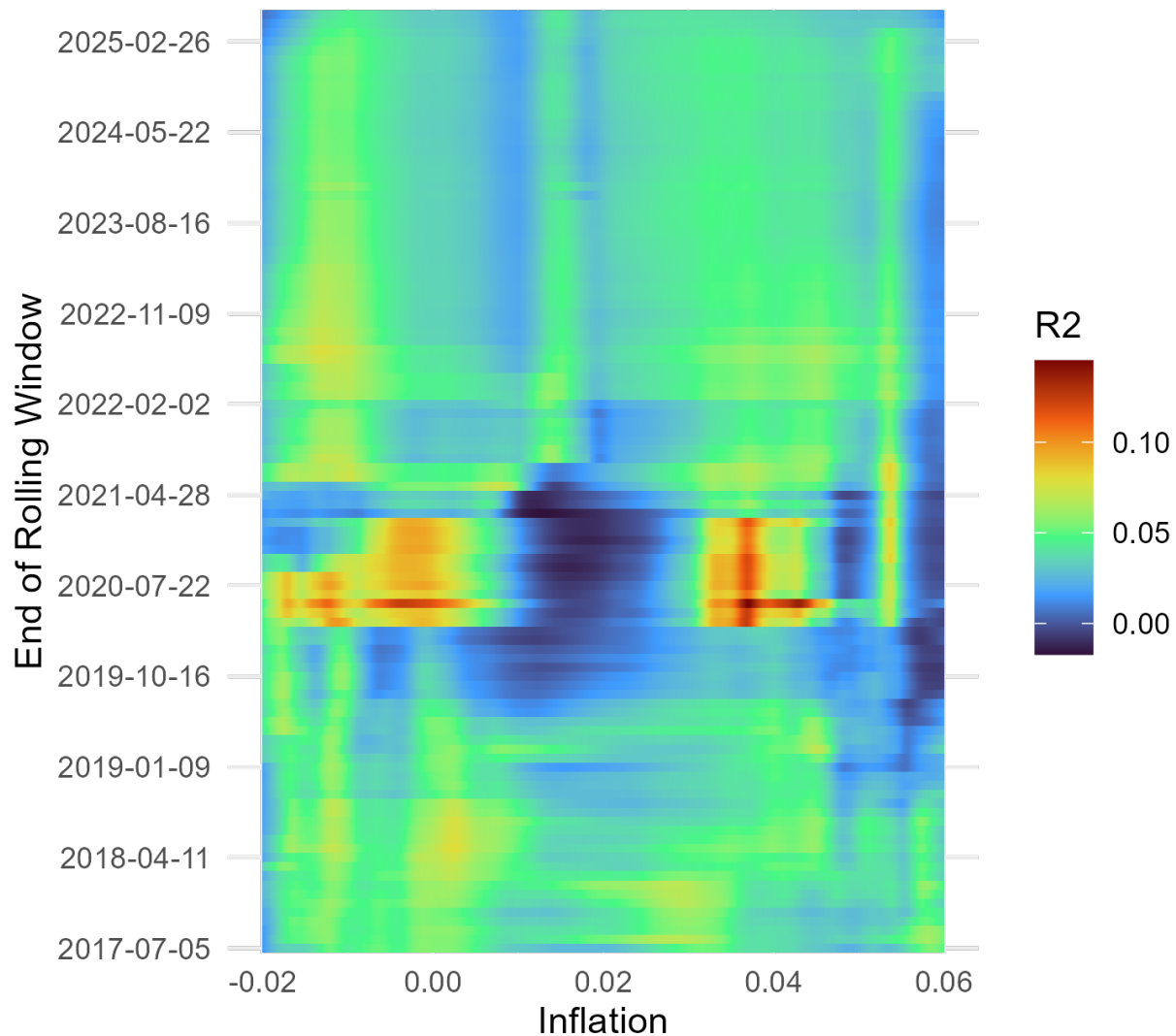
The conditional R^2 values are substantially smaller than the total R^2 in Figure 5.6, indicating that macroeconomic news explains little of the deanchoring. Two features stand out:

Unconditional unanchoring dominates. Comparing Figures 5.6 and 5.7, the vast majority of explained variation comes from the intercept β_1 , not macro surprises. That is, the unconditional R^2 —persistent deviations from $\mu^*(s)$ —dominates the conditional R^2 from news sensitivity. This means violations of unconditional anchoring (Definition 5.3.2)—persistent shifts in the expectations density away from $\mu^*(s)$ —drive the patterns in Figure 5.6, especially around 2020. The expectations density and the policymaker’s target density are not well-aligned unconditionally.

Modest conditional unanchoring during pandemic. There is, however, evidence that for roughly one year during the pandemic 2020–2021, macroeconomic news variables provide some explanatory power in both the left and right tails, with conditional R^2 values reaching 0.10–0.15. This indicates temporary violations of conditional anchoring (Definition 5.3.3) during the pandemic period, though the effect is smaller and more transient than the unconditional shifts.

support as the end of the 5-year windows approach 2018.

Figure 5.7: Conditional functional $R^2(s)$ excluding the intercept – to isolate the variation explained by the macroeconomic surprises (Definition 5.3.3).

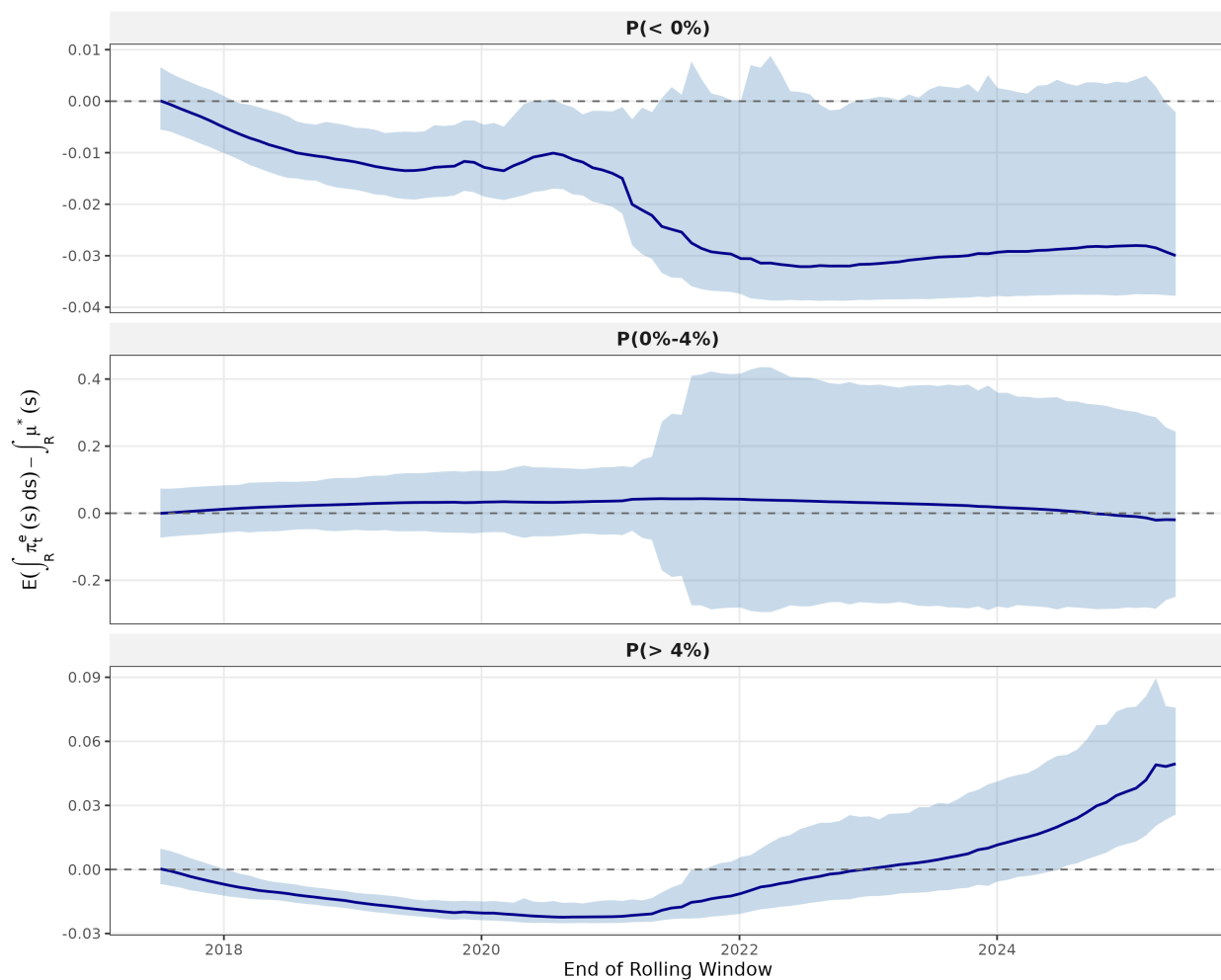


Good versus bad regional anchoring

Figures 5.6 and 5.7 show that expectations deviated from target and where on the distribution, but do not reveal whether these deviations represent heightened or diminished disaster risk. Figure 5.8 addresses this by testing regional anchoring (Definition 5.3.6) across three regions: the tolerance region $R_\epsilon = [0\%, 4\%]$, the lower tail ($s < 0\%$), and the upper tail ($s > 4\%$). In this test, we compute the cumulative differences in probability mass within each region between our predicted densities and the μ^* , and we also produce 95 percent

credible intervals around our predicted densities. The credible intervals allow us to formally test for statistical differences between the predicted densities and μ^* , so that if the credible interval includes zero then we reject regional anchoring.

Figure 5.8: Testing unconditional regional anchoring across three regions: left tail, center, upper tail.



Notes: Regional differences in cumulative probabilities between predicted densities and the policy mean function. Shaded bands are 95% credible intervals implied by the functional coefficients. When the interval excludes zero, we reject regional anchoring. Credible intervals computed using HAC-adjusted posterior covariance.

2018–2020: Good regional anchoring. The rise in global R^2 through late-2020 reflects *good* regional anchoring. The top panel shows the expected probability of inflation being less than 0% is statistically significantly lower than the policymaker’s target $\alpha^* = \int_{R_\epsilon} \mu^*(s) ds$.

The bottom panel tells a similar story for the expected probability of inflation being greater than 4%. These results suggest that the market concentrated less probability below 0% and above 4% than $\mu^*(s)$ prescribes. This withdrawal of probability from disaster regions indicates lower tail risk than the policymaker’s benchmark—a benign violation of perfect anchoring (Definition 5.3.4).

2021–2025: Bad upper-tail anchoring. From late-2021, we observe a marked increase in uncertainty (widening credible intervals). This coincides with the start of the post-pandemic inflation surge. But despite monetary policy being tightened in response through 2022 and 2023, by late-2024, we find statistically significant evidence of *bad upper-tail anchoring*: excess probability mass resides in the region $s > 4\%$, indicating heightened risk of persistent high inflation. The probability of inflation exceeding 4% rose from approximately 5% (consistent with $\mu^*(s)$) to 12% by 2025. Critically, this upper-tail unanchoring occurred even as mean expectations remained near 2%—a pattern that would be invisible to traditional point-forecast metrics like those of Kumar et al. (2015) or Gürkaynak et al. (2010). Bad upper-tail anchoring explains the rise in global R^2 seen in Figure 5.5 from 2024.

Several interpretations are consistent with persistent bad upper-tail anchoring despite disinflation from late-2023. First, it may reflect genuine uncertainty from financial market participants about whether disinflation will persist or prove temporary. The post-pandemic inflation surge may have durably raised perceived tail risks, even if central expectations recovered. Second, it could reflect elevated inflation risk premia: investors may demand greater compensation for inflation risk following the 2021-2022 experience, shifting the RND (Q-distribution) away from real-world beliefs (P-distribution). Our data do not allow us to separate these interpretations, as RNDs conflate probability beliefs and risk premia.

5.5.3 Labor Market Effects on Upper-Tail Unanchoring

Having established *when* (Fact 1) and *where* (Fact 2) probabilistic anchoring failed, we now examine which macroeconomic surprises drive these patterns. We do this by inspecting the back-transformed functional coefficients $\beta_j(s)$ from equation (5.21) over two specific windows: 2016–2020 (good regional anchoring) and 2021–2025 (bad upper-tail anchoring).¹⁸

2016–2020: Good anchoring driven by unconditional shift

Figure 5.9 plots the density-space coefficient functions for the 2016–2020 period. The intercept $\beta_1(s)$ is the only consistently statistically significant coefficient. Its functional form shows how probability mass was reallocated *away* from both tails and *toward* the center of the distribution: $\beta_1(s) > e(s) = 12.5$ (the uniform density) for $s \in [1\%, 3\%]$, and $\beta_1(s) < 12.5$ in both tails. This confirms the good central-region probabilistic anchoring documented in Figure 5.8.

In contrast, the coefficients on macroeconomic news variables show relatively smaller deviations from the uniform distribution. Because surprise measures are standardized, we can compare magnitudes directly. The fed funds rate coefficient shows the largest conditional effect: a positive surprise (actual rate exceeds forecasts) reallocates probability mass away from the left side of the distribution toward the right tail. This pattern is consistent with market participants viewing unexpected FOMC tightening as signaling higher future inflation, or with changes in inflation risk premia demanded by investors hedging policy uncertainty. Similar patterns at smaller magnitudes appear for GDP and PPI surprises.

Notably, nonfarm payrolls shows the opposite pattern: positive employment surprises shift mass slightly toward the left tail. This likely reflects a singular outlier—a 20-standard-deviation surprise in early 2020 (see Figure 5.3) coinciding with pandemic-related leftward shifts in the RNDs. This highlights the reduced-form nature of our analysis and the need

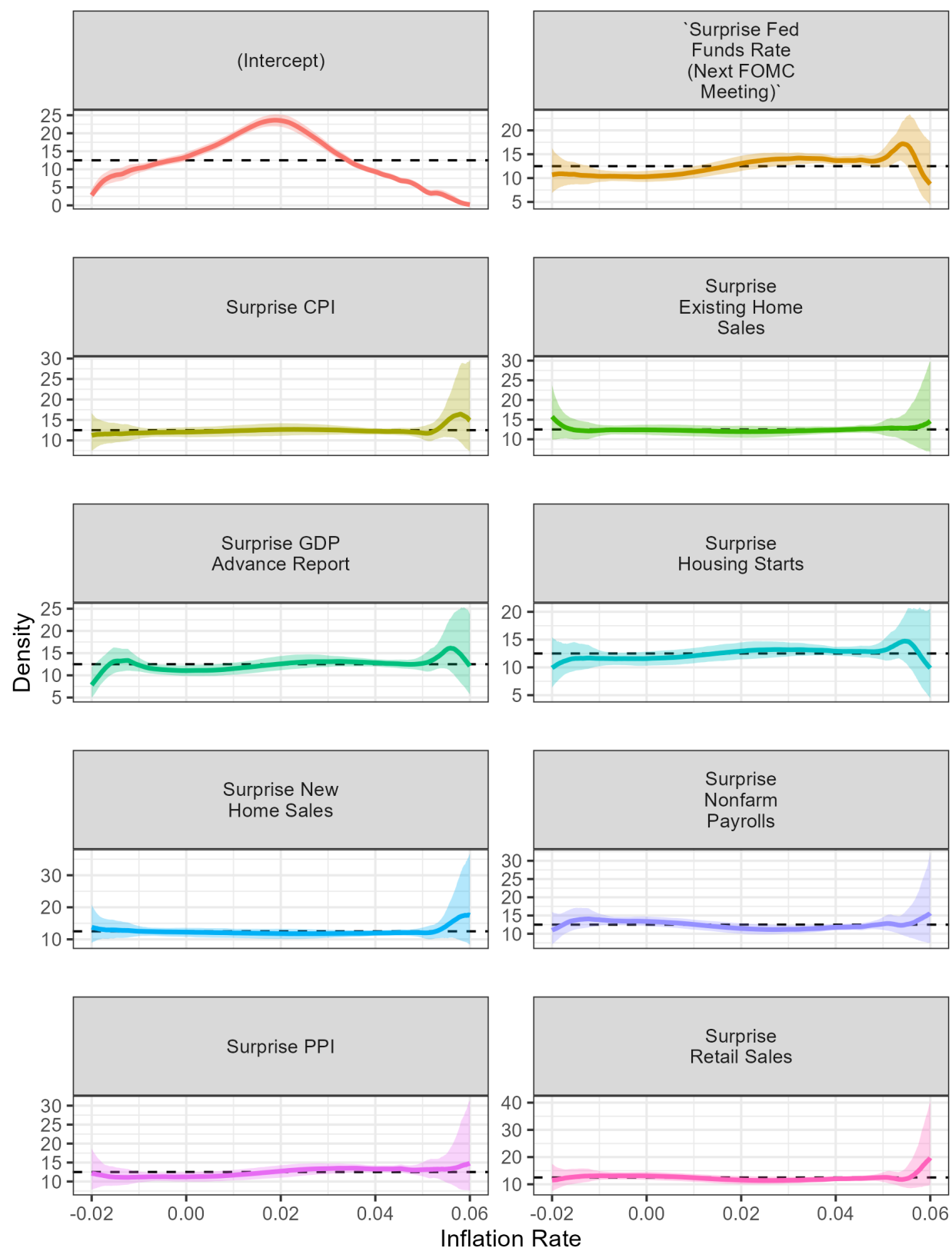
¹⁸Given the reduced form nature of our analysis, we do not claim these effects to be causal. That being said, we do mitigate any impacts of endogeneity bias by the lagged structure of our macroeconomic surprises relative to the risk-neutral densities in our data.

for caution in any causal interpretation.¹⁹

2021–2025: Nonfarm payrolls drives conditional unanchoring

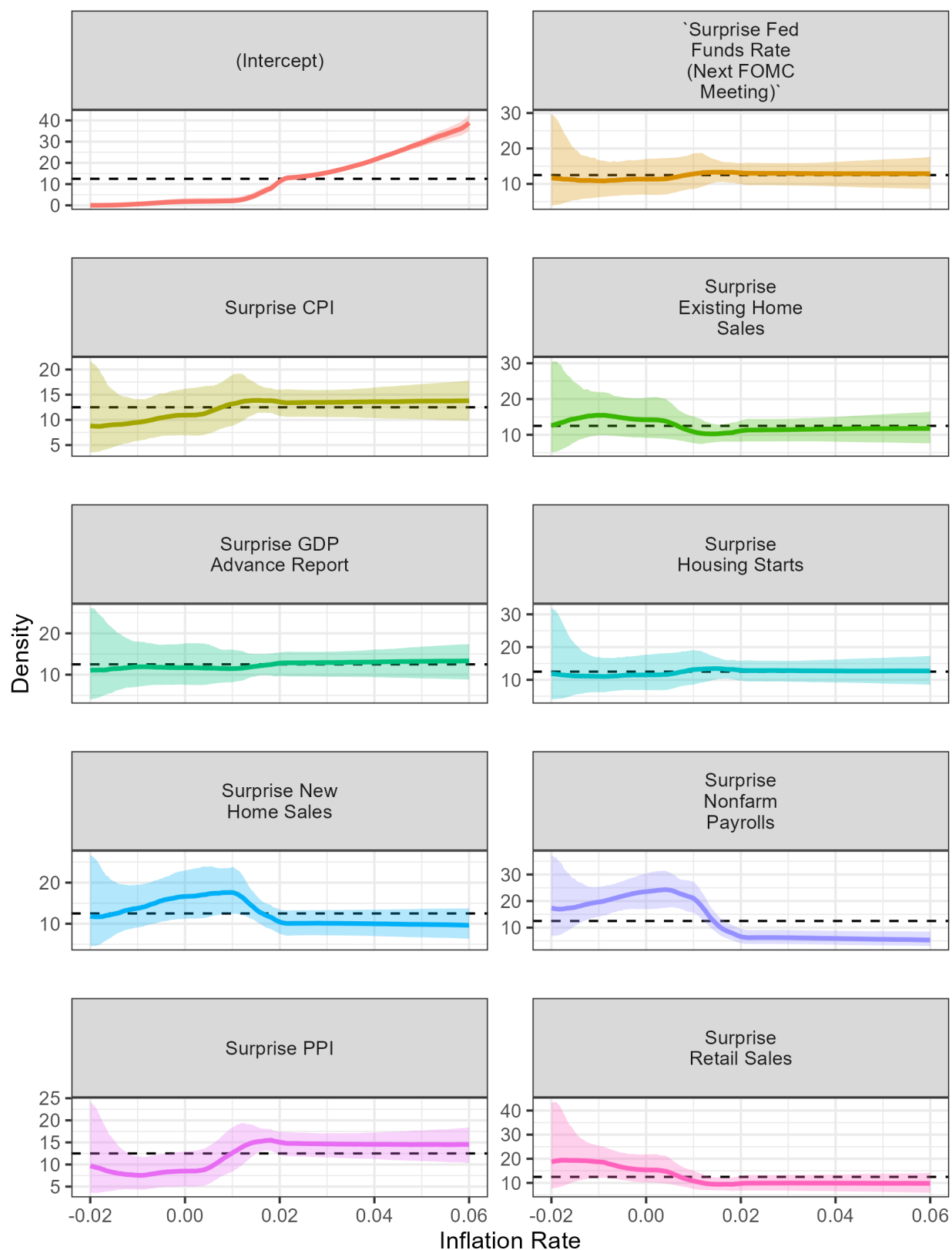
Figure 5.10 shows the functional coefficients for 2021–2025. The pattern differs starkly from 2016–2020:

¹⁹Partial R^2 decompositions in Figure 5.E confirm the intercept dominates, explaining nearly 70% of total variation near 2% inflation. Among macro surprises, the fed funds rate explains the most (5–6.5% of variation in the [3%, 5.5%] region), while other surprises explain negligible variation.... Figure 5.F zooms in on the partial R^2 of the macroeconomic surprises.

Figure 5.9: Functional coefficients $\beta_j(s)$ for 2016–2020, back-transformed to density space

Notes: Shaded regions are 95% credible intervals constructed using HAC-adjusted posterior covariance.

Macroeconomic surprises are standardized, so magnitudes are directly comparable. Dashed line represents the neutral density $e(s) = 12.5$. Coefficients above (below) this line indicate reallocation of probability mass toward (away from) inflation level s .

Figure 5.10: Functional coefficients $\beta_j(s)$ for 2021–2025, back-transformed to density space

Notes: Shaded regions are 95% credible intervals constructed using HAC-adjusted posterior covariance. Macroeconomic surprises are standardized, so magnitudes are directly comparable. Dashed line represents the neutral density $e(s) = 12.5$. Coefficients above (below) this line indicate reallocation of probability mass toward (away from) inflation level s .

Unconditional rightward shift. The most notable feature is the residual intercept $\beta_1(s)$, which now shows a reallocation of probability mass *away* from the left tail and *toward* the right tail—consistent with the bad upper-tail anchoring documented in Figure 5.8. Specifically, $\beta_1(s) < 12.5$ for $s < 1\%$ and $\beta_1(s) > 12.5$ for $s > 3\%$, indicating systematic withdrawal of probability from deflation scenarios and accumulation in high-inflation scenarios. This unconditional shift explains why the probability of inflation exceeding 4% rose from 5% to 12% even as mean expectations remained near 2%.

Nonfarm payrolls drives conditional unanchoring. The next largest functional coefficient in terms of both magnitude and statistical significance is nonfarm payrolls. Intriguingly, the coefficient appears counter-intuitive: positive employment surprises (actual > forecast) reallocate probability mass *away* from the region above 2% toward the region below. This pattern—where strong labor market news appears to lower inflation expectations—likely reflects the specific macroeconomic context of 2021–2025. During this period, unexpectedly strong employment growth may have signaled either: (i) greater economic resilience reducing stagflation fears, or (ii) increased Federal Reserve credibility in achieving a “soft landing” without persistent inflation. Other surprises show similar patterns but are smaller in magnitude or lack statistical significance.

Partial R^2 decompositions (Figure 5.H in the appendix) confirm these interpretations. The residual intercept has the largest explanatory power in the far tails, particularly for $s > 4\%$ where it explains 25–35% of total variation. Among macroeconomic surprises, nonfarm payrolls is the dominant driver, explaining 5–10% of variation in the $[0\%, 4\%]$ region—nearly an order of magnitude larger than other surprises (Figure 5.I).

Implications. The prominence of labor market surprises in explaining conditional unanchoring has two implications. First, it suggests that the market’s inflation expectations are sensitive to employment news in ways that may reflect beliefs about the Phillips curve trade-off or the Fed’s dual mandate. When labor markets surprise to the upside, agents apparently

revise *downward* their probability of extreme inflation outcomes, perhaps interpreting strong employment as signaling policy credibility. Second, this sensitivity indicates that Federal Reserve communication about the employment-inflation nexus may be crucial for maintaining probabilistic anchoring. If the public systematically misinterprets labor market strength as inflationary (or disinflationary), communication clarifying the policy reaction function could help re-anchor tail probabilities.

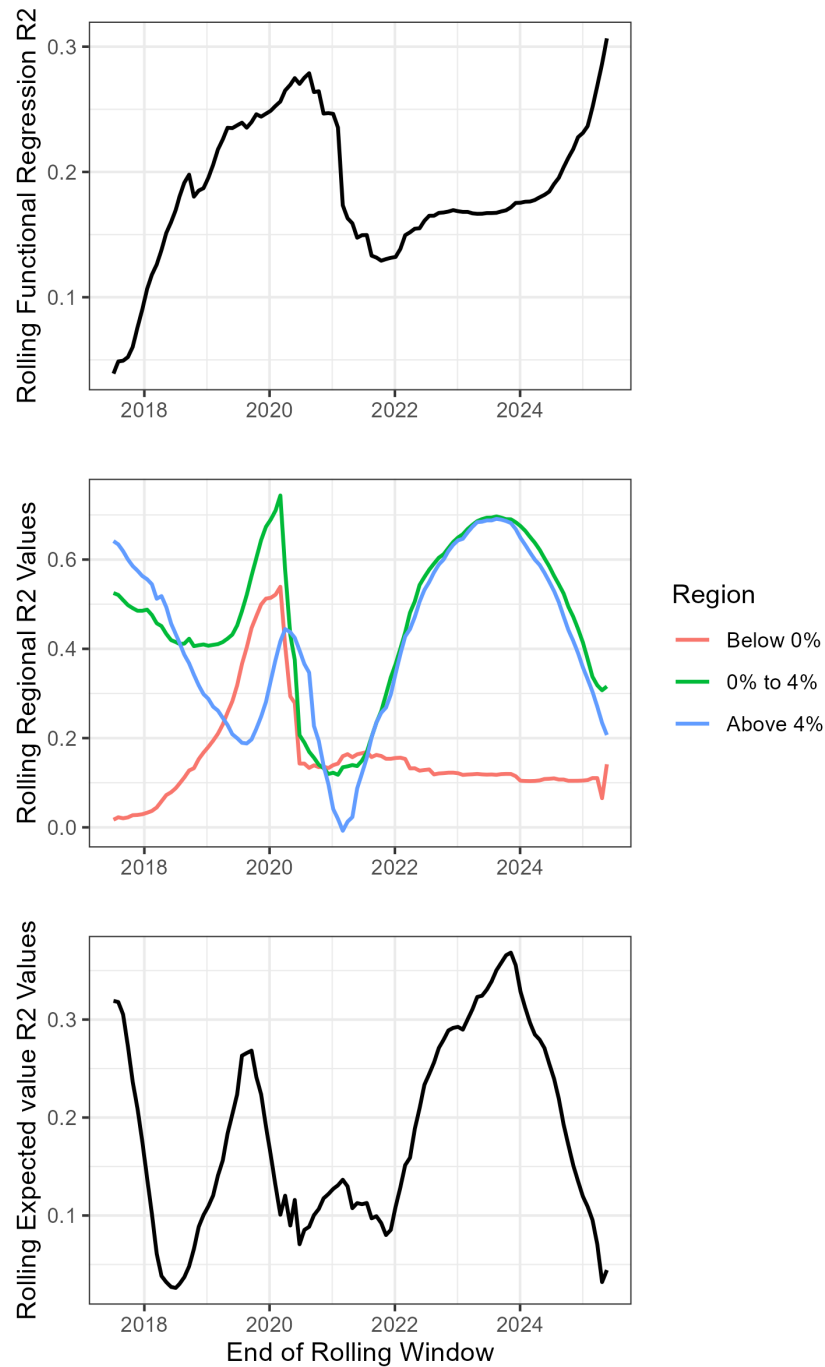
5.5.4 Comparison to Traditional Metrics

We end by comparing our probabilistic tests to traditional anchoring tests. We estimate two benchmark specifications on the same rolling windows as used for the functional regressions:

1. **Mean regression (Kumar et al. variant):** OLS regression of $(\bar{\pi}_t^e - 2)$ on an intercept and macro surprises, where $\bar{\pi}_t^e = \int s \cdot \pi_t^e(s) ds$ is the mean expectation.
2. **Regional probability regression (Grishchenko et al. variant):** OLS regression of $[\alpha_t - \alpha^*]$ on an intercept and macro surprises, where $\alpha_t = \Pr(\pi_t^e \in [0\%, 4\%])$.

Figure 5.11 plots the rolling R^2 from these specifications alongside our global R^2 from Figure 5.5.

Figure 5.11: Comparison of rolling R^2 across three approaches: our distributional method (top panel), regional probability regression (middle panel), and mean regression (bottom panel).



These comparisons demonstrate that traditional point-forecast metrics can mask substantial distributional unanchoring. Our global R^2 peaks at 0.30 in 2025, capturing the full extent

of distributional shifts that expected value and regional probability metrics understate. For example, during late 2025 when the RNDs had begun to place increasing tail mass at higher levels of inflation, our functional global R^2 suggests that the functional approach is able to capture the tail migration while the regional and expected value R^2 metrics suggest largely declining variation explained during that time. Moreover, only our functional coefficients reveal *where* on the distribution news matters (center versus tails) and *which* surprises drive unanchoring. By contrast, Table 5.B shows that the non-zero effects of the macroeconomic surprises on the expectations density would be missed if one simply regressed them on the probability of inflation being less than zero, between zero and four, and greater than four.

5.6 Conclusions

We develop and implement a probabilistic measure of inflation expectations anchoring that extends traditional point-based metrics to assess whether the entire distribution of beliefs aligns with policymakers' targets. Using Bayesian function-on-scalar regression in Bayes space, we test whether risk-neutral densities from inflation options markets are unconditionally consistent with a target density and conditionally orthogonal to macroeconomic surprises across the full support.

Our empirical findings reveal three key findings about probabilistic anchoring over 2012-2025. First, total anchoring deteriorated in two distinct waves (2018-2020 and 2024-2025), with the global R^2 rising from near zero to 0.30. Second, the nature of unanchoring fundamentally changed: the 2018-2020 wave reflected good regional anchoring (excess concentration near 2%), while the 2024-2025 wave reflects bad upper-tail anchoring (12% probability of inflation exceeding 4%, up from 5%). Third, labor market surprises, particularly nonfarm payrolls, drive upper-tail unanchoring in the recent period—a finding invisible to traditional point-forecast metrics that focus solely on mean expectations.

These findings have important implications for monetary policy. First, they demonstrate

that distributional anchoring can fail even when point expectations appear well-behaved, suggesting that central banks should always monitor the full distribution of inflation expectations, not just mean or median forecasts. Second, the persistence of bad upper-tail anchoring through 2025—despite aggressive policy tightening and actual disinflation—suggests distributional anchoring exhibits hysteresis. Once tail beliefs become unanchored, restoring confidence may require sustained periods of low inflation, not just immediate policy responses. Third, the sensitivity of tail probabilities to labor market news indicates that communication about the Fed’s employment mandate and the Phillips curve trade-off may affect perceived inflation risks. Finally, our framework provides policymakers with tools to assess not just whether expectations are anchored on average, but where on the distribution anchoring may be weakening—enabling more targeted policy responses. Future work could combine our analysis of market-based risk-neutral densities with survey-based density forecasts from firms, professional forecasters, and households. Examining heterogeneity in probabilistic anchoring across different types of agents—e.g., comparing market-based measures (reflecting sophisticated investors) with household survey expectations—could reveal whether distributional unanchoring occurs differentially across the population.

Table 5.A: Summary statistics of macroeconomic surprises, observations from July 2012-May 2025. Source: Action Economics, Authors' calculations.

Variable	N	Mean	SD	Min	Median	Max
CPI	675	0.000	0.001	-0.003	0.000	0.005
Existing Home Sales	675	0.000	0.078	-0.640	0.000	0.660
Fed Funds Rate (Next FOMC Meeting)	675	0.000	0.000	-0.005	0.000	0.002
GDP Advance Report	675	0.000	0.002	-0.019	0.000	0.015
Housing Starts	675	0.000	0.038	-0.187	0.000	0.248
New Home Sales	675	0.001	0.024	-0.169	0.000	0.191
Nonfarm Payrolls	675	8.009	209.197	-837.000	0.000	4509.000
PPI	675	0.000	0.001	-0.008	0.000	0.012
Retail Sales	675	0.000	0.004	-0.029	0.000	0.082

Online Appendix for Chapter 5

Additional Tables

Table 5.B: Regional Anchoring Regressions.

	p(< 0)	p(0-4)	p(> 4)
(Intercept)	-0.011*** (0.001)	-0.001 (0.007)	0.012 (0.007)
Surprise CPI	0.000 (0.001)	0.005 (0.003)	-0.006 (0.004)
Surprise Existing Home Sales	-0.000 (0.000)	-0.003 (0.002)	0.004 (0.003)
Surprise Housing Starts	0.001 (0.001)	-0.000 (0.003)	-0.001 (0.003)
Surprise New Home Sales	0.000 (0.000)	0.000 (0.002)	-0.001 (0.002)
Surprise Nonfarm Payrolls	-0.001 (0.002)	-0.013 (0.008)	0.014 (0.009)
Surprise PPI	0.000 (0.000)	0.003 (0.002)	-0.003* (0.002)
Surprise Retail Sales	0.000 (0.001)	0.004 (0.004)	-0.004 (0.005)
Surprise GDP Advance Report	-0.001 (0.000)	-0.003 (0.002)	0.003 (0.002)
Surprise Fed Funds Rate (Next FOMC Meeting)	0.001 (0.001)	0.002 (0.002)	-0.002 (0.003)
Num.Obs.	104	104	104
R2	0.056	0.069	0.074
RMSE	0.01	0.03	0.03

* p < 0.1, ** p < 0.05, *** p < 0.01

Dependent variable is the difference between the cumulative probability implied by the RNDs and the policy mean function at selected intervals of inflation. Newey-West correction used for standard errors.

5.7 Additional Figures

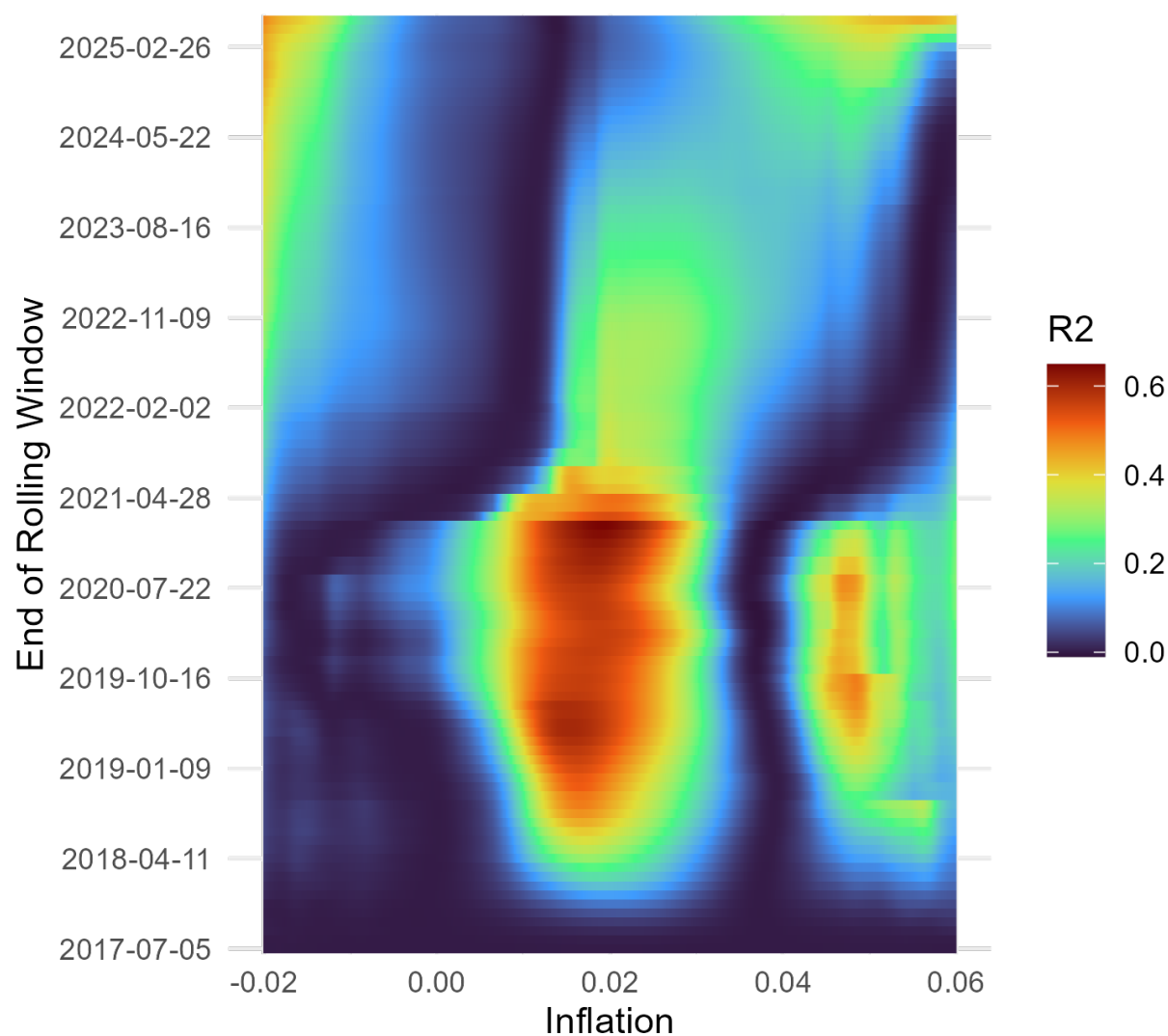


Figure 5.A: Rolling R^2 heatmap, excluding macroeconomic surprises. Functional regressions are estimated at a monthly frequency over a rolling 5-year window of data beginning in the period from July 2012–July 2017.

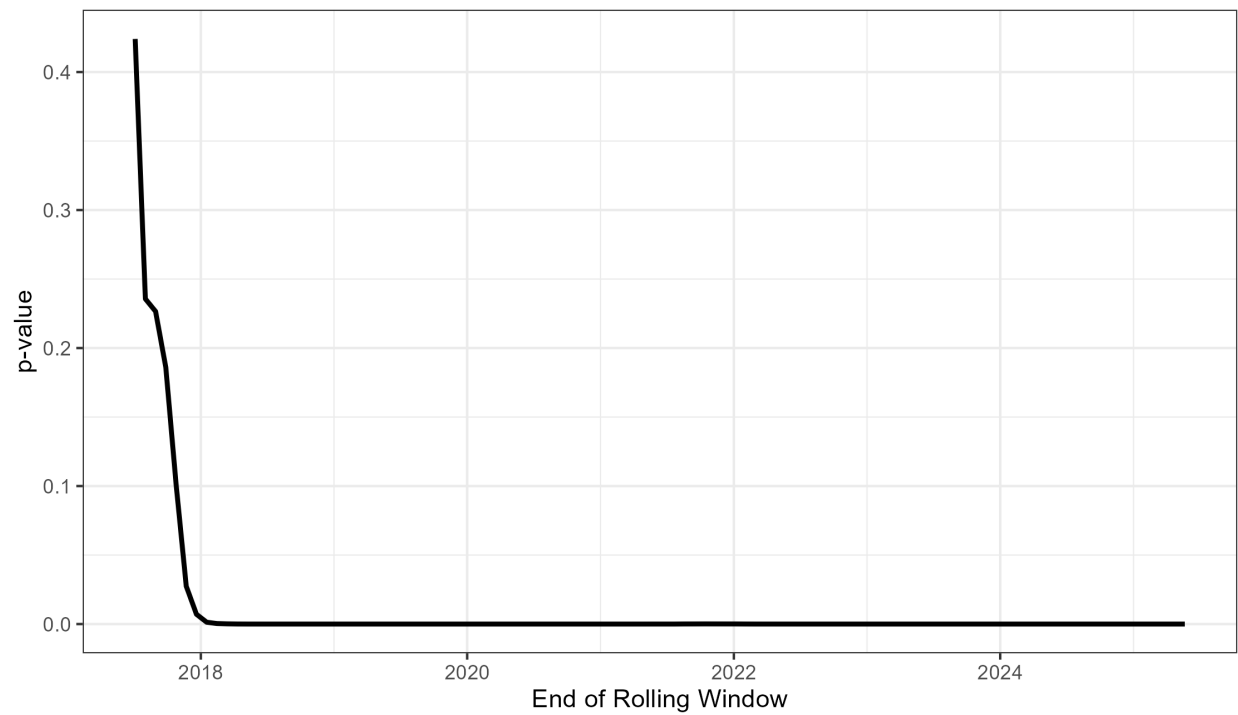


Figure 5.B: Rolling global F-test p-values. Functional regressions are estimated at a monthly frequency over a rolling 5-year window of data beginning in the period from July 2012-July 2017.

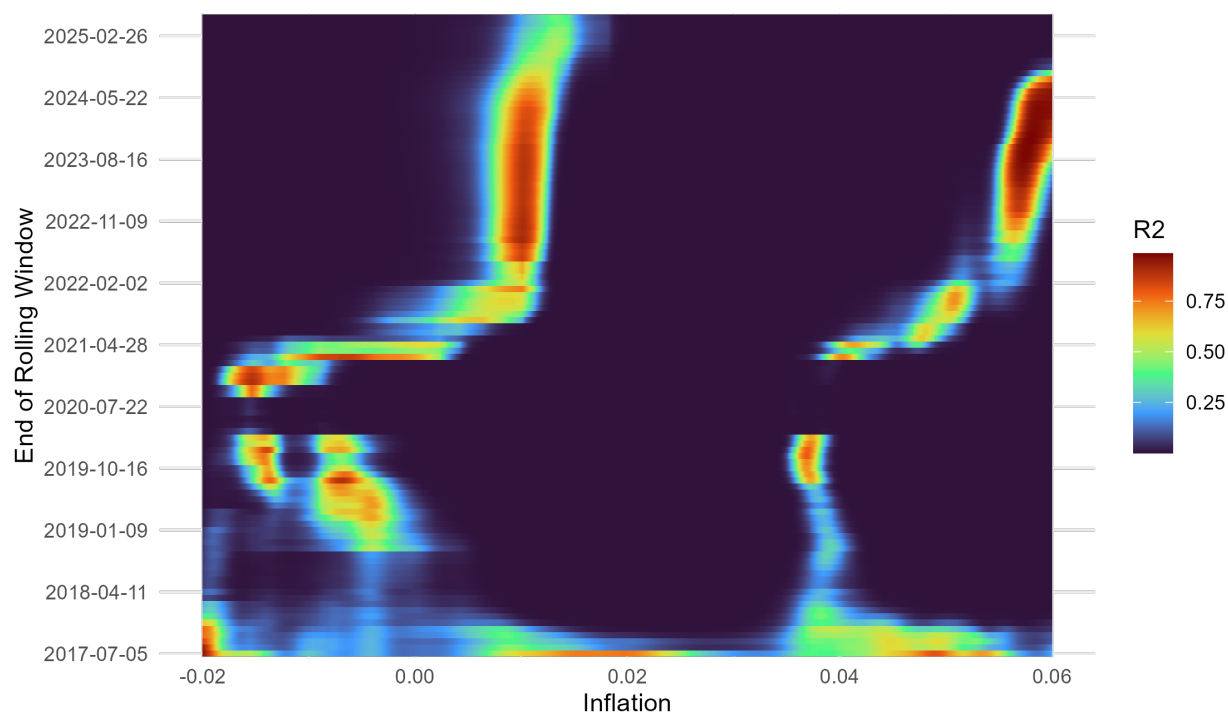


Figure 5.C: Rolling F-test p-value heatmap. Functional regressions are estimated at a monthly frequency over a rolling 5-year window of data beginning in the period from July 2012-July 2017.

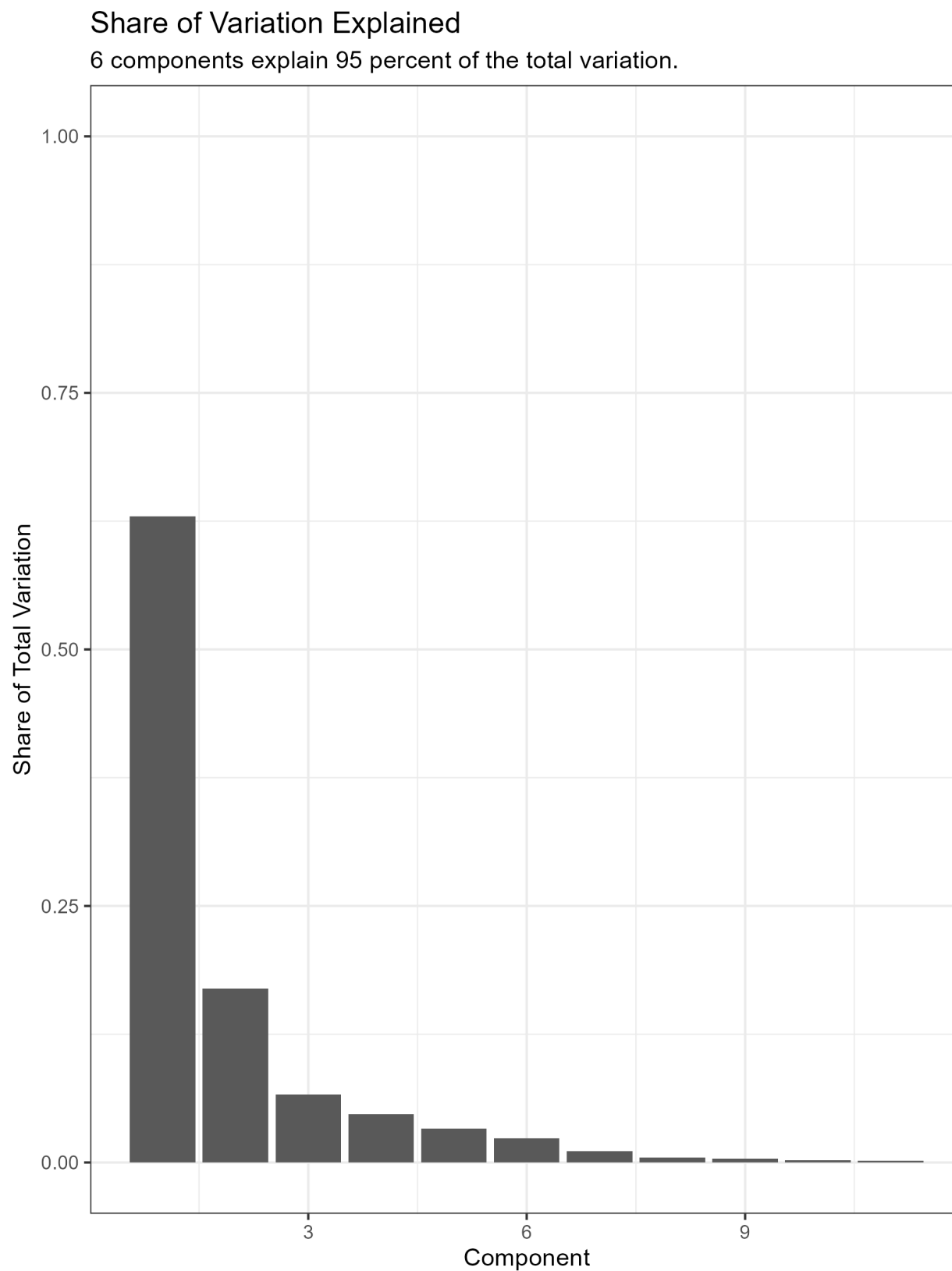


Figure 5.D: Share of total variation explained by functional principal components in CLR space over the period from 2016-2020.

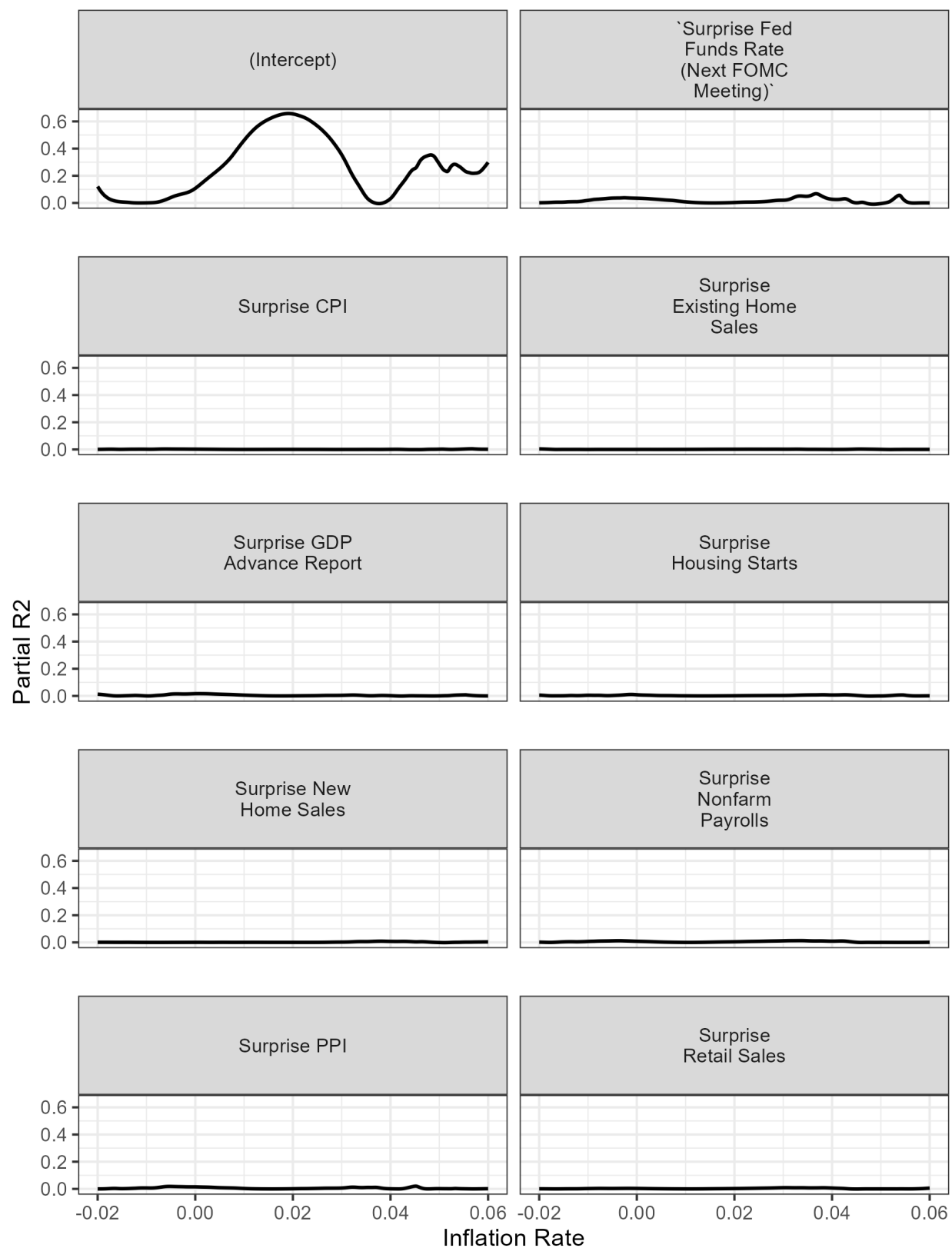


Figure 5.E: Density space partial- R^2 functions using density and macro surprise data from 2016-2020.

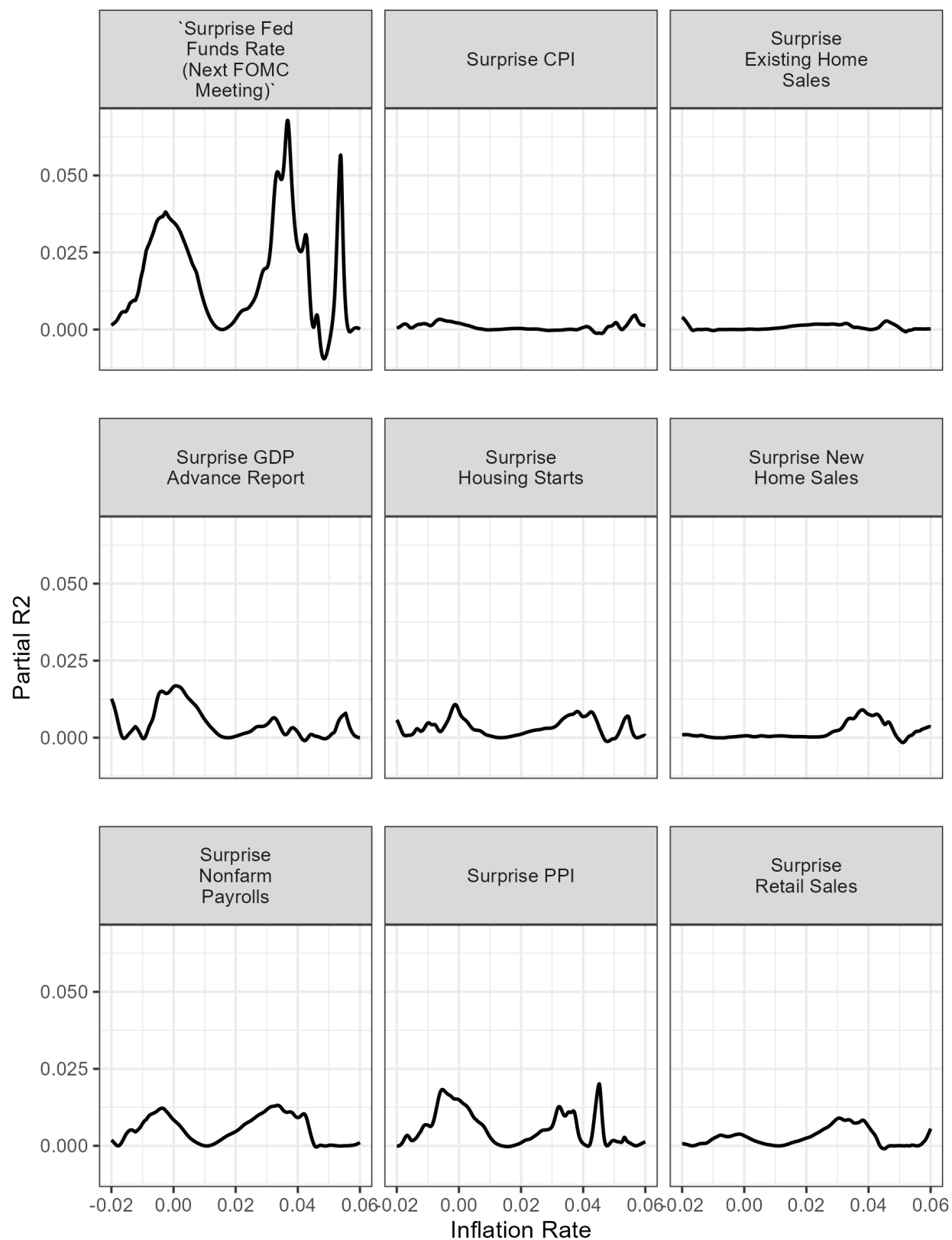


Figure 5.F: Density space partial- R^2 functions, excluding the intercept, using density and macro surprise data from 2016-2020.

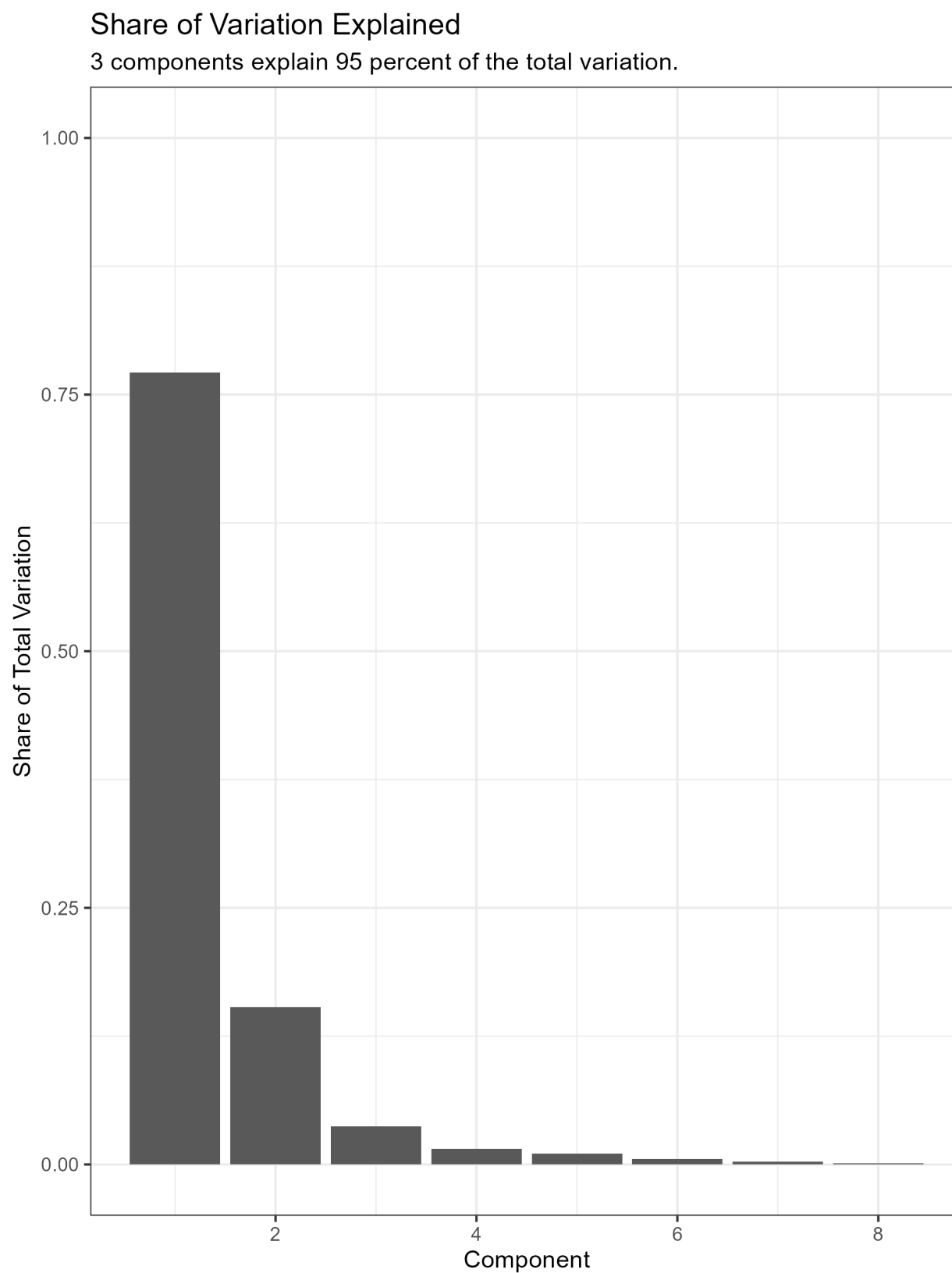


Figure 5.G: Share of total variation explained by functional principal components in CLR space over the period from 2021-2025.

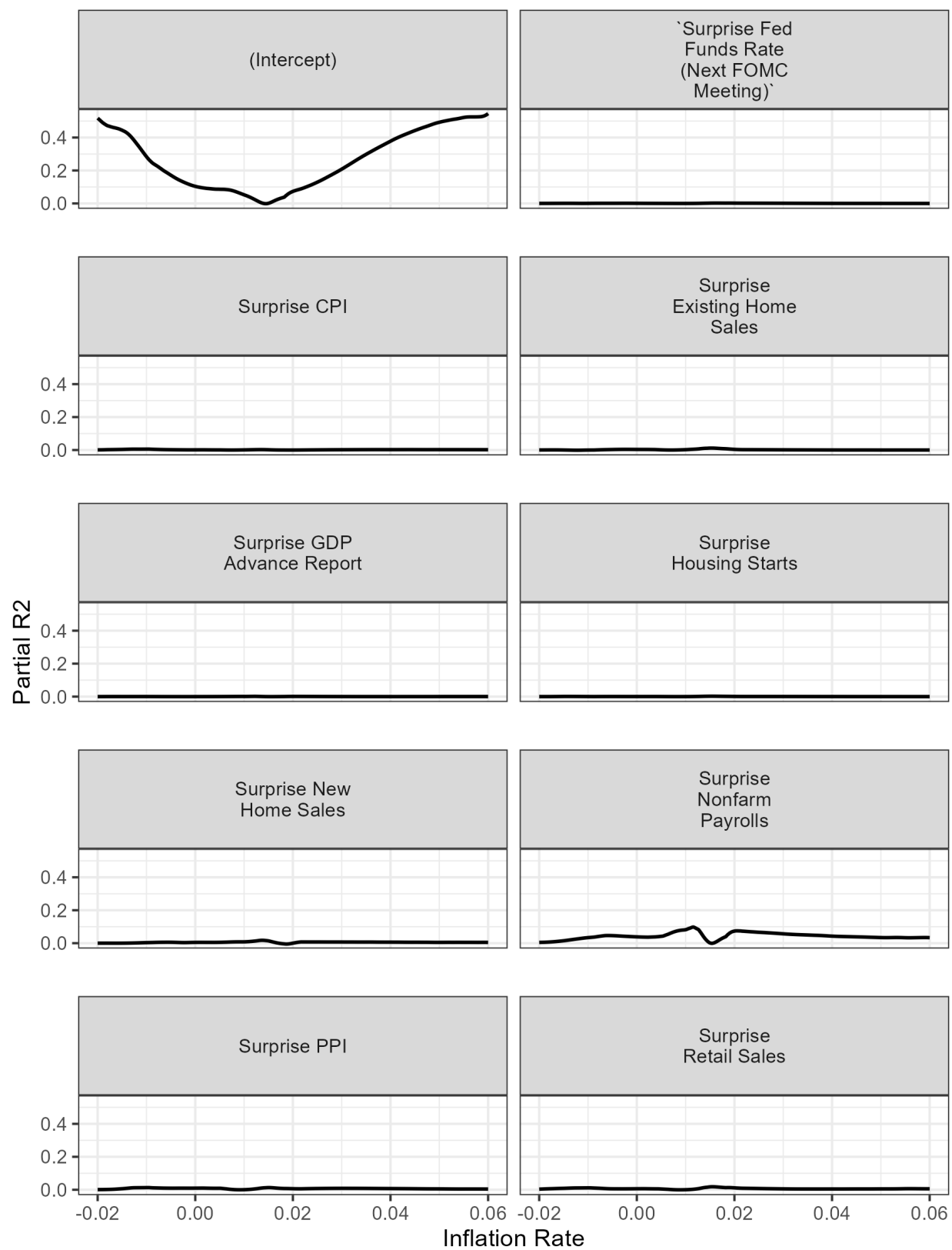


Figure 5.H: Density space partial- R^2 functions using density and macro surprise data from 2021-2025.

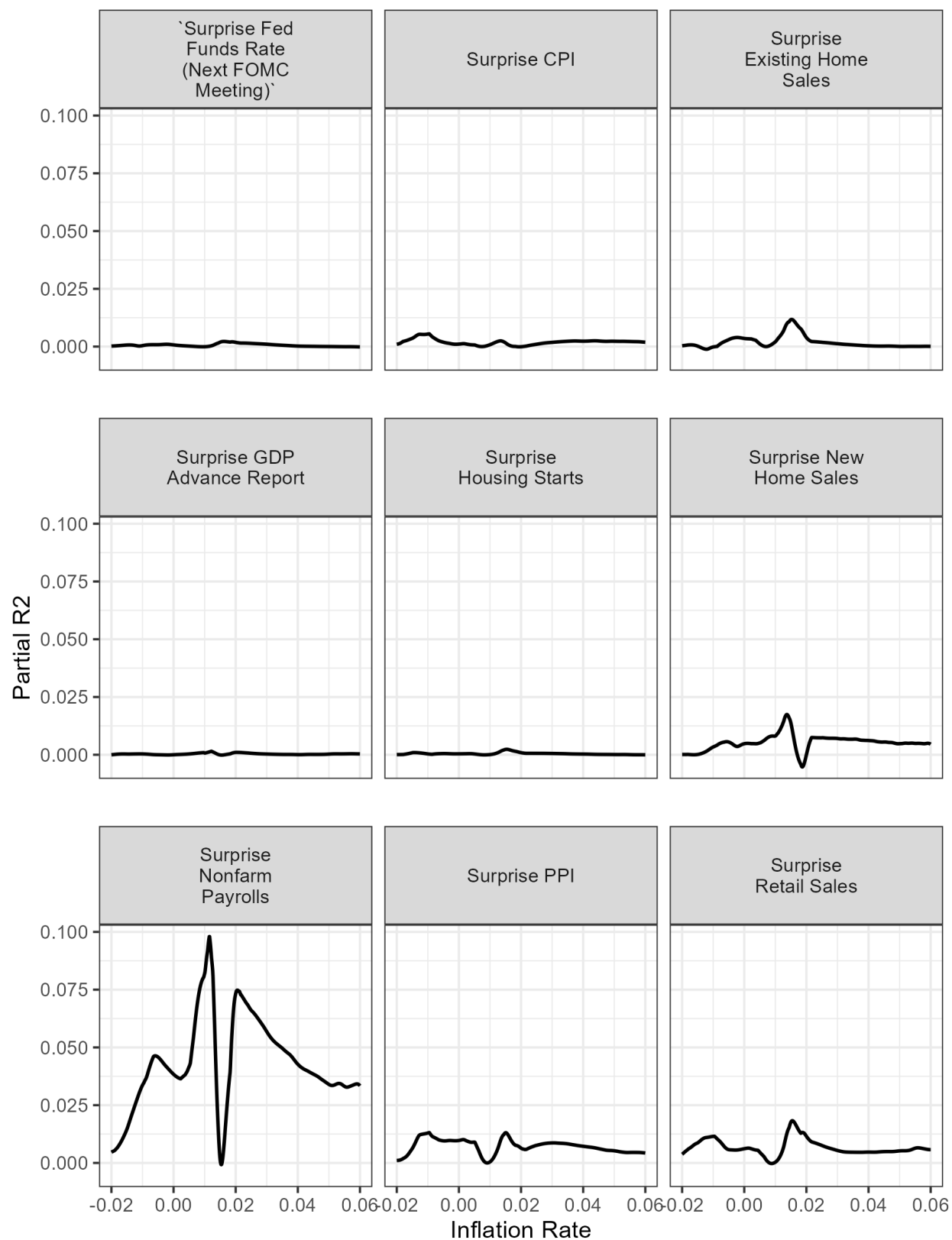


Figure 5.I: Density space partial- R^2 functions, excluding the intercept, using density and macro surprise data from 2021-2025.

Chapter 6

Conclusion

6.1 Summary of Findings

This thesis has examined three questions in the areas of empirical banking and finance: the transmission of interest rate risk through bank profitability, the stability of credit risk models in consumer lending, and measurement of anchoring of inflation expectations. Each essay has documented evidence of significant instability or heterogeneity in relationships that are commonly assumed to be stable in both academic research and policy applications.

Chapter 1 examined the relationship between net interest margins and interest rates, documenting heterogeneity in how banks' profitability responds to interest rate changes. The analysis showed increased sensitivity to interest rates in the tails of the conditional NIM distribution during and after the pandemic period, with nonlinearities resulting from bank interest expenses constrained by the zero lower bound. I document additional heterogeneity by bank size, as smaller community banks experienced different margin dynamics than larger banks that did not reprice deposits upward as quickly prior to the pandemic. These findings have implications for financial stability given banks rely on earnings to build loss-absorbing capital during periods of stress.

Chapter 2 demonstrated that the sensitivity of auto loan default probabilities to unem-

ployment shocks has declined persistently following both the Global Financial Crisis and the Covid recession. The estimated effect of a one percentage point increase in unemployment on default risk fell from 16% in 2006 to just 3% in the post-2020 period. This decline persists even after controlling for forbearance programs and pandemic-specific factors, and is driven primarily by changes in macroeconomic relationships rather than changes in borrower composition. Models estimated just before the pandemic would have overstated two-year cumulative defaults by 100 basis points (25%) during 2020-2021, with implications for stress testing practices and loss forecasting by lenders and regulators.

Chapter 3 developed a probabilistic measure of inflation expectations anchoring and applied it to risk-neutral densities from inflation options over 2012-2025. The analysis revealed that anchoring deteriorated in two waves with fundamentally different characteristics. The 2018-2020 wave reflected excess concentration of probability mass very close to the 2% target (“good” regional anchoring), while the 2024-2025 wave reflected elevated probabilities in the upper tail of the distribution (“bad” tail unanchoring). This upper-tail unanchoring is correlated with labor market surprises and would be completely missed by traditional metrics that focus only on point expectations.

6.2 Common Themes

Two notable common themes emerge across these three essays that have broader implications for empirical research and policy analysis.

6.2.1 Empirical Instability

A central finding is that relationships which appear stable during normal times can break down during periods of extreme stress or large policy interventions. Bank margins responded differently to the rapid tightening of 2022-2023 than to previous periods in the tails of the distributions. The unemployment-default relationship weakened dramatically after both the

GFC and Covid recessions. Inflation expectations became unanchored in their distributional characteristics during and after the pandemic period of 2020.

This evidence of empirical instability poses challenges for risk management and policy models that are based on reduced form approaches. Interest rate transmission through the banking system depends on stable margin responses to rate changes, yet these responses vary across time and institutions. Stress testing models, by their nature, aim to predict outcomes during tail events—precisely when historical relationships may no longer hold. Monetary policy rules and forecasting models assume stable expectation formation processes—yet the anchoring of inflation expectations appears to vary over time and across the distribution.

The Lucas Critique has association with these findings. If economic relationships depend on the policy regime, then using reduced-form statistical models estimated under one regime to evaluate policies in another regime is fundamentally problematic. The massive fiscal and monetary interventions during Covid, the rapid tightening of 2022-2023, and the various forbearance programs implemented during the pandemic all represent regime changes that altered the structural relationships underlying our empirical models.

6.2.2 The Importance of Distributional Analysis

A second theme is the value of examining full distributions rather than focusing exclusively on conditional means or point forecasts. Chapter 1’s examination of heterogeneity across bank types reveals that aggregate margin responses mask important differences in how monetary policy affects different segments of the banking system. Chapter 2’s Markov transition framework captures the full dynamics of loan performance—delinquency, cure, prepayment, and default—rather than just terminal default rates. This approach allows for more accurate multi-period projections.

Chapter 3’s distributional approach to inflation expectations reveals unanchoring in the tails of the distribution that is invisible in mean expectations. If central banks monitor only whether mean five-year inflation expectations remain near 2%, they will miss important

changes in perceived tail risks. Yet it is precisely these tail risks—the probability of inflation running persistently above target or falling into deflation—that matter most for policy.

This emphasis on distributions connects to a broader trend in macroeconomics and finance toward examining risk, uncertainty, and higher moments of distributions rather than just central tendencies. The Global Financial Crisis highlighted the importance of tail risks that appeared small in normal times. The growing availability of density forecast data from both surveys and financial markets has enabled researchers to move beyond point expectations.

6.3 Policy Implications

The findings in this thesis suggest several implications for policy and risk management practice.

The evidence of heterogeneous net interest margin responses to interest rate changes has important implications for understanding interest rate transmission through the banking system. Loss-absorbing bank capital is built from retained earnings, so it is crucial for policymakers to understand how changes in both policy rates and the broader yield curve may impact the tails of the bank margin distribution when measuring risks to bank solvency, a core component of financial stability. If bank earnings and resulting solvency are negatively impacted by interest rate changes during a crisis then banks could reduce lending to businesses and households, leading to broad effects on the real economy.

The evidence of declining unemployment sensitivity in auto loan defaults raises questions about the stability of the macroeconomic relationships embedded in stress testing models. If models calibrated during earlier periods systematically overstate credit losses during recessions, this could lead to inefficiently high capital requirements. Banks may be required to hold excess capital to guard against losses that are less likely to materialize than historical relationships suggest. One possible solution is to update model parameters more frequently

to reflect recent data. An alternative is to simply acknowledge this uncertainty explicitly by incorporating sensitivity analysis around key parameters when forecasting losses.

The finding that inflation expectations can become unanchored in their distributional characteristics while mean expectations remain stable has important implications for how central banks monitor and communicate about expectations. It is common for central banks to focus heavily on whether mean long-run inflation expectations remain near 2%. Our results suggest that central banks should expand their monitoring frameworks to include distributional measures of anchoring, utilizing statistical tests for distributional anchoring. When distributional anchoring deteriorates—as it did in 2024-2025—policymakers should recognize this as a warning sign even if mean expectations appear stable.

6.4 Concluding Remarks

This thesis has examined three important questions in empirical banking, risk management and inflation expectations, documenting evidence of significant instability and heterogeneity in relationships that are central to policy models and risk management frameworks. Bank margins respond heterogeneously to monetary policy changes. The sensitivity of auto loan defaults to unemployment has declined persistently. Inflation expectations can become unanchored in their distributional characteristics while mean expectations remain stable.

These findings highlight challenges and opportunities facing empirical researchers and policymakers. The challenge is that relationships we thought we understood appear to be changing, making prediction and policy design more difficult. But the opportunity is to develop richer empirical approaches—examining full distributions, testing for heterogeneity, and explicitly modeling instability—that provide a more complete picture of economic relationships and their evolution over time. Econometric methods already exist to deal with these issues, but it is not clear that they are fully utilized in current policy frameworks. This thesis has augmented an existing literature highlighting the need for greater use of these

techniques.

The period from 2000 to 2025 has been extraordinary in terms of macroeconomic volatility and policy experimentation. The Global Financial Crisis, the Covid pandemic and policy response, and the subsequent inflation surge and rapid tightening have each represented natural experiments that have revealed the limits of models estimated during more stable periods. As we move forward, the lessons from this period should inform how we build and evaluate empirical models, conduct policy analysis, and manage risk. In addition, it could be that appropriately calibrated structural models would offer more stability in their ability to predict across structural changes in the economy. Future research could explore how structural models may outperform the reduced form models typically used in the empirical banking, stress testing, and inflation literature while continuing to focus on heterogeneity and distributional analysis as opposed to more common representative agent assumptions.

Bibliography

- Abdymomunov, Azamat, Jeffery Gerlach, and Yuji Sakurai (2023). “Interest Rate Risk in the U.S. Banking Sector”. en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: [10.2139/ssrn.4395529](https://doi.org/10.2139/ssrn.4395529). URL: <https://www.ssrn.com/abstract=4395529> (visited on 03/25/2024).
- Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone (Apr. 2019). “Vulnerable Growth”. en. In: *American Economic Review* 109.4, pp. 1263–1289. ISSN: 0002-8282. DOI: [10.1257/aer.20161923](https://doi.org/10.1257/aer.20161923). URL: <https://pubs.aeaweb.org/doi/10.1257/aer.20161923> (visited on 09/01/2023).
- Agarwal, Sumit (2008). “Determinants of Automobile Loan Default and Prepayment;” en. In: *Economic Perspectives* 32.3.
- Agarwal, Sumit and Brent W. Ambrose, eds. (2007). *Household Credit Usage*. en. New York: Palgrave Macmillan US. ISBN: 978-1-349-53946-8 978-0-230-60891-7. DOI: [10.1057/9780230608917](https://doi.org/10.1057/9780230608917). URL: <http://link.springer.com/10.1057/9780230608917> (visited on 02/17/2026).
- Aghamohammadi, A. and S. Mohammadi (Dec. 2017). “Bayesian analysis of penalized quantile regression for longitudinal data”. en. In: *Statistical Papers* 58.4, pp. 1035–1053. ISSN: 0932-5026, 1613-9798. DOI: [10.1007/s00362-015-0737-4](https://doi.org/10.1007/s00362-015-0737-4). URL: <http://link.springer.com/10.1007/s00362-015-0737-4> (visited on 03/25/2024).
- Alessandri, Piergiorgio and Benjamin D. Nelson (Feb. 2015). “Simple Banking: Profitability and the Yield Curve”. en. In: *Journal of Money, Credit and Banking* 47.1, pp. 143–175.

- ISSN: 00222879. DOI: [10.1111/jmcb.12172](https://doi.org/10.1111/jmcb.12172). URL: <https://onlinelibrary.wiley.com/doi/10.1111/jmcb.12172> (visited on 09/01/2023).
- Alhamzawi, Rahim, Keming Yu, and Dries F Benoit (2012). “Bayesian adaptive Lasso quantile regression”. In: *Statistical Modelling* 12.3, pp. 279–297. DOI: [10.1177/1471082X1101200304](https://doi.org/10.1177/1471082X1101200304).
- An, Xudong, Larry Cordell, and Sharon Tang (May 2020). *Extended Loan Terms and Auto Loan Default Risk*. en. Working paper (Federal Reserve Bank of Philadelphia) 20-18. Series: Working paper (Federal Reserve Bank of Philadelphia). Federal Reserve Bank of Philadelphia, pp. 20–18. DOI: [10.21799/frbp.wp.2020.18](https://doi.org/10.21799/frbp.wp.2020.18). URL: <https://www.philadelphiafed.org/-/media/frbp/assets/working-papers/2020/wp20-18.pdf> (visited on 02/16/2026).
- Angbazo, Lazarus (Jan. 1997). “Commercial bank net interest margins, default risk, interest-rate risk, and off-balance sheet banking”. en. In: *Journal of Banking & Finance* 21.1, pp. 55–87. ISSN: 03784266. DOI: [10.1016/S0378-4266\(96\)00025-8](https://doi.org/10.1016/S0378-4266(96)00025-8). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0378426696000258> (visited on 09/01/2023).
- Azzalini, Adelchi and Antonella Capitanio (May 2003). “Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t -Distribution”. en. In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 65.2, pp. 367–389. ISSN: 1369-7412, 1467-9868. DOI: [10.1111/1467-9868.00391](https://doi.org/10.1111/1467-9868.00391). URL: <https://academic.oup.com/jrsssb/article/65/2/367/7092856> (visited on 09/01/2023).
- Bakshi, Tanya and Rose, Jonathan (2021). “What happened to subprime auto loans during the COVID-19 pandemic”. In: *Chicago Fed Insights*.
- Begenau, Juliane and Erik Stafford (2022). “Unstable Inference from Banks’ Stable Net Interest Margins”. In: *Available at SSRN* 4136866. DOI: [10.2139/ssrn.4136866](https://doi.org/10.2139/ssrn.4136866).
- Bems, Rudolfs, Francesca Caselli, Francesco Grigoli, and Bertrand Gruss (2021). “Expectations’ anchoring and inflation persistence”. In: *Journal of International Economics* 132, p. 103516. ISSN: 0022-1996. DOI: [10.1016/j.jinteco.2021.103516](https://doi.org/10.1016/j.jinteco.2021.103516).

- Bernanke, Ben S. (Feb. 2004). *The Great Moderation*. Remarks at the meetings of the Eastern Economic Association, Washington, DC. URL: <https://www.federalreserve.gov/boarddocs/speeches/2004/20040220/>.
- Bernanke, Ben S. (July 2007). *Inflation expectations and inflation forecasting*. Speech at the NBER Summer Institute, Cambridge, Massachusetts. URL: <https://www.federalreserve.gov/newsevents/speech/bernanke20070710a.htm>.
- Bianchi, Francesco, Leonardo Melosi, and Matthias Rottner (2021). “Hitting the elusive inflation target”. In: *Journal of Monetary Economics* 124, pp. 107–122. ISSN: 0304-3932. DOI: [10.1016/j.jmoneco.2021.10.003](https://doi.org/10.1016/j.jmoneco.2021.10.003).
- Boero, Gianna, Jeremy Smith, and Kenneth F. Wallis (2008). “Uncertainty and Disagreement in Economic Prediction: The Bank of England Survey of External Forecasters”. In: *Economic Journal* 118.530, pp. 1107–1127. DOI: [10.1111/j.1468-0297.2008.02162.x](https://doi.org/10.1111/j.1468-0297.2008.02162.x).
- Borio, Claudio, Leonardo Gambacorta, and Boris Hofmann (2017). “The influence of monetary policy on bank profitability”. en. In: *International Finance* 20.1. _eprint: <https://onlinelibrary.wiley.com/doi/abs/10.1111/infi.12104> (visited on 04/01/2024). pp. 48–63. ISSN: 1468-2362. DOI: [10.1111/infi.12104](https://doi.org/10.1111/infi.12104). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/infi.12104> (visited on 04/01/2024).
- Breeden, Douglas T and Robert H Litzenberger (1978). “Prices of state-contingent claims implicit in option prices”. In: *Journal of business*, pp. 621–651.
- Bundick, Brent and A. Lee Smith (2025). “Did the Federal Reserve Break the Phillips Curve? Theory and Evidence of Anchoring Inflation Expectations”. In: *Review of Economics and Statistics* 107.5, pp. 1310–1326. DOI: [10.1162/rest_a_01357](https://doi.org/10.1162/rest_a_01357).
- Busetti, Fabio (Aug. 2017). “Quantile Aggregation of Density Forecasts”. en. In: *Oxford Bulletin of Economics and Statistics* 79.4, pp. 495–512. ISSN: 0305-9049, 1468-0084. DOI: [10.1111/obes.12163](https://doi.org/10.1111/obes.12163). URL: <https://onlinelibrary.wiley.com/doi/10.1111/obes.12163> (visited on 03/25/2024).

- Canals-Cerda, José J. and Brian Jonghwan Lee (2021). “COVID-19 and Auto Loan Origination Trends”. en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: [10.2139/ssrn.3897908](https://doi.org/10.2139/ssrn.3897908). URL: <https://www.ssrn.com/abstract=3897908> (visited on 02/16/2026).
- Capistrán, Carlos (2008). “Bias in Federal Reserve inflation forecasts: Is the Federal Reserve irrational or just cautious?” In: *Journal of Monetary Economics* 55.8, pp. 1415–1427. ISSN: 0304-3932. DOI: [10.1016/j.jmoneco.2008.09.011](https://doi.org/10.1016/j.jmoneco.2008.09.011).
- Capistrán, Carlos and Manuel Ramos-Francia (2010). “Does Inflation Targeting Affect the Dispersion of Inflation Expectations?” In: *Journal of Money, Credit and Banking* 42.1, pp. 113–134. DOI: [10.1111/j.1538-4616.2009.00280.x](https://doi.org/10.1111/j.1538-4616.2009.00280.x).
- Carriero, Andrea, Todd E Clark, and Massimiliano Marcellino (2020). *Nowcasting Tail Risks to Economic Activity with Many Indicators*. Tech. rep. Federal Reserve Bank of Cleveland. DOI: [10.26509/frbc-wp-202013r2](https://doi.org/10.26509/frbc-wp-202013r2).
- Carriero, Andrea, Todd E Clark, and Massimiliano Marcellino (2024). “Specification choices in quantile regression for empirical macroeconomics”. In: *Journal of Applied Econometrics*. DOI: [10.1002/jae.3099](https://doi.org/10.1002/jae.3099).
- Carvalho, Carlos, Stefano Eusepi, Emanuel Moench, and Bruce Preston (Jan. 2023). “Anchored Inflation Expectations”. In: *American Economic Journal: Macroeconomics* 15.1, pp. 1–47. DOI: [10.1257/mac.20200080](https://doi.org/10.1257/mac.20200080).
- Chang, Minsu, Xiaohong Chen, and Frank Schorfheide (2024). “Heterogeneity and Aggregate Fluctuations”. In: *Journal of Political Economy* 132.6, pp. 2045–2106. DOI: [10.1086/729444](https://doi.org/10.1086/729444).
- Chen, Qingqing, Dennis Glennon, and Amos Golan (2020). “Estimating Conditional Mortgage Delinquency Transition Matrices”. en. In: *OCC Working Paper Series*.
- Clarida, Richard, Jordi Galí, and Mark Gertler (2000). “Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory”. In: *Quarterly Journal of Economics* 115.1. Also available as NBER Working Paper 6442, pp. 147–180. DOI: [10.1162/003355300554692](https://doi.org/10.1162/003355300554692).

- Clements, Michael P. (2014). “Forecast Uncertainty—Ex Ante and Ex Post: U.S. Inflation and Output Growth”. In: *Journal of Business & Economic Statistics* 32.2, pp. 206–216. DOI: [10.1080/07350015.2013.859618](https://doi.org/10.1080/07350015.2013.859618).
- Clements, Michael P. and Shixuan Wang (2024). *Do Professionals’ Inflation Forecasts Incorporate the Beliefs of Others? A Functional Data Approach*. Working Paper 4357012. SSRN. DOI: [10.2139/ssrn.4357012](https://doi.org/10.2139/ssrn.4357012).
- Cline, Alexander and Robert W. Rich (Jan. 2026). *How Anchored Are Short-Run Inflation Expectations Today? A Look at What Consumers and Forecasters Are Telling Us*. Economic Commentary 2026-01. Federal Reserve Bank of Cleveland. DOI: [10.26509/frbc-ec-202601](https://doi.org/10.26509/frbc-ec-202601).
- Covas, Francisco, Ben Rump, and Egon Zakrajšek (July 2014). “Stress-testing US bank holding companies: A dynamic panel quantile regression approach”. en. In: *International Journal of Forecasting* 30.3, pp. 691–713. ISSN: 01692070. DOI: [10.1016/j.ijforecast.2013.11.003](https://doi.org/10.1016/j.ijforecast.2013.11.003). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0169207013001623> (visited on 09/01/2023).
- Deng, Yongheng, John M. Quigley, and Robert Order (Mar. 2000). “Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options”. en. In: *Econometrica* 68.2, pp. 275–307. ISSN: 0012-9682, 1468-0262. DOI: [10.1111/1468-0262.00110](https://doi.org/10.1111/1468-0262.00110). URL: <http://doi.wiley.com/10.1111/1468-0262.00110> (visited on 02/17/2026).
- Detting, Lisa and Lauren Lambie-Hanson (Mar. 2021). “Why is the Default Rate So Low? How Economic Conditions and Public Policies Have Shaped Mortgage and Auto Delinquencies During the COVID-19 Pandemic”. en. In: *FEDS Notes* 2021.0.2854. ISSN: 2380-7172. DOI: [10.17016/2380-7172.2854](https://doi.org/10.17016/2380-7172.2854). URL: <https://www.federalreserve.gov/econres/notes/feds-notes/why-is-the-default-rate-so-low-20210304.htm> (visited on 02/17/2026).

- Di Lucido, Katherine, Anna Kovner, and Samantha Zeller (June 2017). *Low Interest Rates and Bank Profits*. en-US. URL: <https://libertystreeteconomics.newyorkfed.org/2017/06/low-interest-rates-and-bank-profits/> (visited on 03/25/2024).
- Douglas Smith, L. and Edward C. Lawrence (Sept. 1995). “Forecasting losses on a liquidating long-term loan portfolio”. en. In: *Journal of Banking & Finance* 19.6, pp. 959–985. ISSN: 03784266. DOI: [10.1016/0378-4266\(94\)00065-B](https://doi.org/10.1016/0378-4266(94)00065-B). URL: <https://linkinghub.elsevier.com/retrieve/pii/037842669400065B> (visited on 02/16/2026).
- Drechsler, Itamar, Alexi Savov, and Philipp Schnabl (June 2021). “Banking on Deposits: Maturity Transformation without Interest Rate Risk”. en. In: *The Journal of Finance* 76.3, pp. 1091–1143. ISSN: 0022-1082, 1540-6261. DOI: [10.1111/jofi.13013](https://doi.org/10.1111/jofi.13013). URL: <https://onlinelibrary.wiley.com/doi/10.1111/jofi.13013> (visited on 09/01/2023).
- Egozcue, Juan José, José Luis Díaz-Barrero, and Vera Pawlowsky-Glahn (2006). “Hilbert space of probability density functions based on Aitchison geometry”. In: *Acta Mathematica Sinica* 22.4, pp. 1175–1182.
- English, William (2002). “Interest rate risk and bank net interest margins”. In: *BIS Quarterly Review*. Publisher: Bank for International Settlements. URL: <https://econpapers.repec.org/article/bisbisqtr/0212g.htm> (visited on 03/25/2024).
- English, William, Skander J. Van Den Heuvel, and Egon Zakrajšek (Oct. 2018). “Interest rate risk and bank equity valuations”. en. In: *Journal of Monetary Economics* 98, pp. 80–97. ISSN: 03043932. DOI: [10.1016/j.jmoneco.2018.04.010](https://doi.org/10.1016/j.jmoneco.2018.04.010). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0304393218302101> (visited on 09/01/2023).
- Ennis, Huberto M and Arantxa Jarque (2021). “Bank Lending in the Time of COVID”. In.
- Fairlie, Robert W. (Nov. 2005). “An extension of the Blinder-Oaxaca decomposition technique to logit and probit models”. en. In: *Journal of Economic and Social Measurement* 30.4, pp. 305–316. ISSN: 0747-9662, 1875-8932. DOI: [10.3233/JEM-2005-0259](https://doi.org/10.3233/JEM-2005-0259). URL: <https://journals.sagepub.com/doi/10.3233/JEM-2005-0259> (visited on 02/17/2026).

- Flannery, Mark (1981). “Market Interest Rates and Commercial Bank Profitability: An Empirical Investigation”. en. In: *The Journal of Finance* 36.5. _eprint: <https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1981.tb01078.x>, pp. 1085–1101. ISSN: 1540-6261. DOI: [10.1111/j.1540-6261.1981.tb01078.x](https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1981.tb01078.x). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1981.tb01078.x> (visited on 03/25/2024).
- Galí, Jordi (2008). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton, NJ: Princeton University Press. ISBN: 978-0-691-13316-4.
- Galvao, Antonio F. (Sept. 2011). “Quantile regression for dynamic panel data with fixed effects”. en. In: *Journal of Econometrics* 164.1, pp. 142–157. ISSN: 03044076. DOI: [10.1016/j.jeconom.2011.02.016](https://linkinghub.elsevier.com/retrieve/pii/S0304407611000443). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0304407611000443> (visited on 09/01/2023).
- Gáti, Laura (2023). “Monetary policy and anchored expectations – An endogenous gain learning model”. In: *Journal of Monetary Economics* 140, S37–S47. DOI: [10.1016/j.jmoneco.2023.06.009](https://doi.org/10.1016/j.jmoneco.2023.06.009).
- Giglio, Carla, Frances Shaw, Nicolas Syrichas, and Giuseppe Cappelletti (Feb. 2021). *Stress-Testing Net Trading Income: The Case of European Banks*. en. ECB Working Paper 2021/2525. URL: <https://papers.ssrn.com/abstract=3797145> (visited on 03/25/2024).
- Gneiting, Tilmann (2011). “Quantiles as optimal point forecasts”. In: *International Journal of Forecasting* 27.2, pp. 197–207. ISSN: 0169-2070. DOI: [10.1016/j.ijforecast.2009.12.015](https://doi.org/10.1016/j.ijforecast.2009.12.015).
- Gneiting, Tilmann and Matthias Katzfuss (2014). “Probabilistic forecasting”. In: *Annual Review of Statistics and Its Application* 1.1, pp. 125–151. DOI: [10.1146/annurev-statistics-062713-085831](https://doi.org/10.1146/annurev-statistics-062713-085831).
- Gneiting, Tilmann and Adrian E Raftery (2007). “Strictly proper scoring rules, prediction, and estimation”. In: *Journal of the American statistical Association* 102.477, pp. 359–378. DOI: [10.1198/016214506000001437](https://doi.org/10.1198/016214506000001437).

- Gneiting, Tilmann and Roopesh Ranjan (2011). “Comparing density forecasts using threshold- and quantile-weighted scoring rules”. In: *Journal of Business & Economic Statistics* 29.3, pp. 411–422. DOI: [10.1198/jbes.2010.08110](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”. In: DOI: [10.2139/ssrn.52400](https://doi.org/10.2139/ssrn.52400).
- Grishchenko, Olesya, Sarah Mouabbi, and Jean-Paul Renne (2019). “Measuring Inflation Anchoring and Uncertainty: A U.S. and Euro Area Comparison”. In: *Journal of Money, Credit and Banking* 51.5, pp. 1053–1096. DOI: [10.1111/jmcb.12622](https://doi.org/10.1111/jmcb.12622).
- Gross, Marco and Javier Población (Feb. 2019). “Implications of Model Uncertainty for Bank Stress Testing”. en. In: *Journal of Financial Services Research* 55.1, pp. 31–58. ISSN: 0920-8550, 1573-0735. DOI: [10.1007/s10693-017-0275-4](https://doi.org/10.1007/s10693-017-0275-4). URL: <http://link.springer.com/10.1007/s10693-017-0275-4> (visited on 09/01/2023).
- Guerrieri, Luca and Michelle Welch (July 2012). *Can Macro Variables Used in Stress Testing Forecast the Performance of Banks?* en. SSRN Scholarly Paper. Rochester, NY. DOI: [10.2139/ssrn.2192650](https://doi.org/10.2139/ssrn.2192650). URL: <https://papers.ssrn.com/abstract=2192650> (visited on 03/25/2024).
- Guo, Zhengfeng, Yan Zhang, and Xinlei Zhao (2022). “Risks of long-term auto loans”. In: *Journal of Credit Risk*. ISSN: 17446619, 17559723. DOI: [10.21314/JCR.2022.009](https://doi.org/10.21314/JCR.2022.009). URL: <https://www.risk.net/journal-of-credit-risk/7954991/risks-of-long-term-auto-loans> (visited on 02/17/2026).
- Gürkaynak, Refet S., Andrew Levin, and Eric Swanson (Dec. 2010). “Does Inflation Targeting Anchor Long-Run Inflation Expectations? Evidence from the U.S., UK, and Sweden”. In: *Journal of the European Economic Association* 8.6, pp. 1208–1242. DOI: [10.1111/j.1542-4774.2010.tb00553.x](https://doi.org/10.1111/j.1542-4774.2010.tb00553.x).
- Hanweck, Gerald A. and Lisa H. Ryu (Jan. 2005). *The Sensitivity of Bank Net Interest Margins and Profitability To Credit, Interest-Rate, and Term-Structure Shocks Across Bank Product Specializations*. en. FDIC Working Paper 05-02. Rochester, NY. DOI: [10.2139/ssrn.844444](https://doi.org/10.2139/ssrn.844444).

- 2139/ssrn.886727. URL: <https://papers.ssrn.com/abstract=886727> (visited on 03/25/2024).
- Haubrich, Joseph G. and Tristan Young (2019). “Trends in the Noninterest Income of Banks”. en. In: *Federal Reserve Bank of Cleveland Economic Commentary* 2019-14. Publisher: Federal Reserve Bank of Cleveland. DOI: [10.26509/frbc-ec-201914](https://doi.org/10.26509/frbc-ec-201914). URL: <https://ideas.repec.org/a/fip/fedcec/00106.html> (visited on 03/25/2024).
- Heitfield, Erik and Tarun Sabarwal (Dec. 2004). “What Drives Default and Prepayment on Subprime Auto Loans?” en. In: *The Journal of Real Estate Finance and Economics* 29.4, pp. 457–477. ISSN: 0895-5638, 1573-045X. DOI: [10.1023/B:REAL.0000044023.02636.e6](https://doi.org/10.1023/B:REAL.0000044023.02636.e6). URL: <https://link.springer.com/10.1023/B:REAL.0000044023.02636.e6> (visited on 02/16/2026).
- Hirtle, Beverly, Anna Kovner, James Vickery, and Meru Bhanot (Aug. 2016). “Assessing financial stability: The Capital and Loss Assessment under Stress Scenarios (CLASS) model”. en. In: *Journal of Banking & Finance* 69, S35–S55. ISSN: 03784266. DOI: [10.1016/j.jbankfin.2015.09.021](https://doi.org/10.1016/j.jbankfin.2015.09.021). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0378426615002940> (visited on 09/01/2023).
- Ho, Thomas S. Y. and Anthony Saunders (Nov. 1981). “The Determinants of Bank Interest Margins: Theory and Empirical Evidence”. en. In: *The Journal of Financial and Quantitative Analysis* 16.4, p. 581. ISSN: 00221090. DOI: [10.2307/2330377](https://doi.org/10.2307/2330377). URL: <https://www.jstor.org/stable/2330377?origin=crossref> (visited on 09/01/2023).
- Hron, Karel, Alessandra Menafoglio, Matthias Templ, Klára Hrzová, and Peter Filzmoser (2016). “Simplicial principal component analysis for density functions in Bayes spaces”. In: *Computational Statistics & Data Analysis* 94, pp. 330–350.
- Huber, Florian, Massimiliano Marcellino, and Tommaso Tornese (2024a). “The Distributional Effects of Economic Uncertainty”. In: *arXiv preprint arXiv:2411.12655*.
- Huber, Florian, Massimiliano Marcellino, and Tommaso Tornese (2024b). “The Distributional Effects of Economic Uncertainty”. In: *arXiv preprint arXiv:2411.12655*.

- Jochmann, Markus, Gary Koop, and Simon M. Potter (2010). “Modeling the dynamics of inflation compensation”. In: *Journal of Empirical Finance* 17.1, pp. 157–167. ISSN: 0927-5398. DOI: [10.1016/j.jempfin.2009.08.002](https://doi.org/10.1016/j.jempfin.2009.08.002).
- Katcher, Bradley, Geng Li, Alvaro Mezza, and Steve Ramos (July 2024). “One Month Longer, One Month Later? Prepayments in the Auto Loan Market”. en. In: *Finance and Economics Discussion Series* 2024-056, pp. 1–66. ISSN: 1936-2854, 2767-3898. DOI: [10.17016/feds.2024.056](https://doi.org/10.17016/feds.2024.056). URL: <https://www.federalreserve.gov/econres/feds/one-month-longer-one-month-later-prepayments-in-the-auto-loan-market.htm> (visited on 02/16/2026).
- Kitsul, Yuriy and Jonathan H Wright (2013). “The economics of options-implied inflation probability density functions”. In: *Journal of Financial Economics* 110.3, pp. 696–711.
- Koenker, Roger (Oct. 2004). “Quantile regression for longitudinal data”. en. In: *Journal of Multivariate Analysis* 91.1, pp. 74–89. ISSN: 0047259X. DOI: [10.1016/j.jmva.2004.05.006](https://doi.org/10.1016/j.jmva.2004.05.006). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0047259X04001113> (visited on 09/01/2023).
- Kozumi, Hideo and Genya Kobayashi (Nov. 2011). “Gibbs sampling methods for Bayesian quantile regression”. In: *Journal of Statistical Computation and Simulation* 81.11. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/00949655.2010.496117>, pp. 1565–1578. ISSN: 0094-9655. DOI: [10.1080/00949655.2010.496117](https://doi.org/10.1080/00949655.2010.496117). URL: <https://doi.org/10.1080/00949655.2010.496117> (visited on 03/25/2024).
- Kumar, Saten, Hassan Afrouzi, Olivier Coibion, and Yuriy Gorodnichenko (2015). “Inflation Targeting Does Not Anchor Inflation Expectations: Evidence from Firms in New Zealand”. In: *Brookings Papers on Economic Activity* 46.2 (Fall), pp. 151–225. URL: <https://ideas.repec.org/a/bin/bpeajo/v46y2015i2015-02p151-225.html>.
- 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”. en. In: *Federal Reserve Bank of Kansas City, Economic Review*. DOI: [10.18651/ER/v109n1LaliberteSengupta](https://doi.org/10.18651/ER/v109n1LaliberteSengupta).
- Lancaster, Tony (1990). “The econometric analysis of transition data”. In:
- Lucas, Robert E. (1976). “Econometric policy evaluation: A critique”. en. In: *Carnegie-Rochester Conference Series on Public Policy* 1, pp. 19–46. ISSN: 01672231. DOI: [10.1016/S0167-2231\(76\)80003-6](https://doi.org/10.1016/S0167-2231(76)80003-6). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0167223176800036> (visited on 02/17/2026).
- Meeks, Raphael and Francesca Monti (2023). “Heterogeneous Beliefs and the Phillips Curve”. In: *Journal of Monetary Economics* 140. Carnegie-Rochester-NYU Conference Series on Public Policy, S86–S104. DOI: [10.1016/j.jmoneco.2023.10.012](https://doi.org/10.1016/j.jmoneco.2023.10.012).
- Mincer, Jacob A and Victor Zarnowitz (1969). “The evaluation of economic forecasts”. In: *Economic forecasts and expectations: Analysis of forecasting behavior and performance*. NBER, pp. 3–46. URL: <https://www.nber.org/system/files/chapters/c1214/c1214.pdf>.
- Müller, Ulrich K (2013). “Risk of Bayesian inference in misspecified models, and the sandwich covariance matrix”. In: *Econometrica* 81.5, pp. 1805–1849.
- Naggert, Kristoph, Robert W. Rich, and Joseph Tracy (July 2023). *The Anchoring of US Inflation Expectations Since 2012*. Economic Commentary 2023-11. Federal Reserve Bank of Cleveland. DOI: [10.26509/frbc-ec-202311](https://doi.org/10.26509/frbc-ec-202311).
- Nautz, Dieter and Till Strohsal (2015). “Are US inflation expectations re-anchored?” In: *Economics Letters* 127, pp. 6–9. ISSN: 0165-1765. DOI: [10.1016/j.econlet.2014.12.023](https://doi.org/10.1016/j.econlet.2014.12.023).
- Newey, Whitney K. and Kenneth D. West (1987). “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”. In: *Econometrica* 55.3, pp. 703–708. ISSN: 00129682, 14680262. URL: <http://www.jstor.org/stable/1913610> (visited on 02/12/2026).

- Nickell, Stephen (1981). “Biases in Dynamic Models with Fixed Effects”. In: *Econometrica* 49.6. Publisher: [Wiley, Econometric Society], pp. 1417–1426. ISSN: 0012-9682. DOI: [10.2307/1911408](https://doi.org/10.2307/1911408). URL: <https://www.jstor.org/stable/1911408> (visited on 03/25/2024).
- Park, Trevor and George Casella (June 2008). “The Bayesian Lasso”. In: *Journal of the American Statistical Association* 103.482. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1198/016214508000000337>. pp. 681–686. ISSN: 0162-1459. DOI: [10.1198/016214508000000337](https://doi.org/10.1198/016214508000000337). URL: <https://doi.org/10.1198/016214508000000337> (visited on 03/25/2024).
- Petersen, Alexander and Hans-Georg Müller (2016). “Functional data analysis for density functions by transformation to a Hilbert space”. In:
- Reis, Ricardo (2021). “Losing the Inflation Anchor”. In: *Brookings Papers on Economic Activity* Fall, pp. 307–379. DOI: [10.1353/eca.2022.0004](https://doi.org/10.1353/eca.2022.0004).
- Samuelson, Paul A (1945). “The effect of interest rate increases on the banking system”. In: *American economic review* 35.1, pp. 16–27. URL: <https://www.jstor.org/stable/1810106>.
- Sengupta, Rajdeep and Fei Xue (2022). “Do Net Interest Margins for Small and Large Banks Vary Differently with Interest Rates?” en. In: *Federal Reserve Bank of Kansas City, Economic Review*. DOI: [10.18651/ER/v107n1SenguptaXue](https://doi.org/10.18651/ER/v107n1SenguptaXue).
- Smith, L. Douglas and Baiqiang Jin (Oct. 2007). “Modeling exposure to losses on automobile leases”. en. In: *Review of Quantitative Finance and Accounting* 29.3, pp. 241–266. ISSN: 0924-865X, 1573-7179. DOI: [10.1007/s11156-007-0032-0](https://doi.org/10.1007/s11156-007-0032-0). URL: <http://link.springer.com/10.1007/s11156-007-0032-0> (visited on 02/16/2026).
- Smith, L. Douglas, Susan M. Sanchez, and Edward C. Lawrence (June 1996). “A Comprehensive Model for Managing Credit Risk on Home Mortgage Portfolios”. en. In: *Decision Sciences* 27.2, pp. 291–317. ISSN: 0011-7315, 1540-5915. DOI: [10.1111/j.1540-5915.1996.tb01719.x](https://doi.org/10.1111/j.1540-5915.1996.tb01719.x). URL: <http://doi.wiley.com/10.1111/j.1540-5915.1996.tb01719.x> (visited on 02/16/2026).

- Talská, Renáta, Alessandra Menafooglio, Jitka Machalová, Karel Hron, and E Fišerová (2018). “Compositional regression with functional response”. In: *Computational Statistics & Data Analysis* 123, pp. 66–85.
- Tibshirani, Robert (Jan. 1996). “Regression Shrinkage and Selection Via the Lasso”. en. In: *Journal of the Royal Statistical Society: Series B Methodology* 58.1, pp. 267–288. ISSN: 00359246. DOI: [10 . 1111 / j . 2517 - 6161 . 1996 . tb02080 . x](https://doi.org/10.1111/j.2517-6161.1996.tb02080.x). URL: [https : / / onlinelibrary . wiley . com / doi / 10 . 1111 / j . 2517 - 6161 . 1996 . tb02080 . x](https://onlinelibrary.wiley.com/doi/10.1111/j.2517-6161.1996.tb02080.x) (visited on 09/01/2023).
- Van den Boogaart, Karl Gerald, Juan José Egozcue, and Vera Pawlowsky-Glahn (2014). “Bayes hilbert spaces”. In: *Australian & New Zealand Journal of Statistics* 56.2, pp. 171–194.
- Williams, Emily (2020). “Heterogeneity in Net-Interest Income Exposure to Interest Rate Risk and Non-Interest Expense Adjustment”. en. In.
- Woodford, Michael (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.
- Wu, Deming, Ming Fang, and Qing Wang (Dec. 2018). “An empirical study of bank stress testing for auto loans”. en. In: *Journal of Financial Stability* 39, pp. 79–89. ISSN: 15723089. DOI: [10 . 1016 / j . jfs . 2018 . 09 . 005](https://doi.org/10.1016/j.jfs.2018.09.005). URL: [https : / / linkinghub . elsevier . com / retrieve / pii / S1572308918303139](https://linkinghub.elsevier.com/retrieve/pii/S1572308918303139) (visited on 02/16/2026).
- Yu, Keming and Rana A. Moyeed (Oct. 2001). “Bayesian quantile regression”. In: *Statistics & Probability Letters* 54.4, pp. 437–447. ISSN: 0167-7152. DOI: [10 . 1016 / S0167 - 7152 \(01\) 00124 - 9](https://doi.org/10.1016/S0167-7152(01)00124-9). URL: [https : / / www . sciencedirect . com / science / article / pii / S0167715201001249](https://www.sciencedirect.com/science/article/pii/S0167715201001249) (visited on 03/25/2024).
- Zarruk, Emilio R and Jeff Madura (1992). “Optimal bank interest margin under capital regulation and deposit insurance”. In: *Journal of Financial and Quantitative analysis* 27.1, pp. 143–149. DOI: [10 . 2307 / 2331303](https://doi.org/10.2307/2331303).

Zou, Hui (Dec. 2006). “The Adaptive Lasso and Its Oracle Properties”. In: *Journal of the American Statistical Association* 101.476. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1198/016214506000000735>. URL: <https://doi.org/10.1198/016214506000000735> (visited on 03/25/2024).