



Trabalho de Conclusão de Curso

**Análise de Estresse Macroeconômico e seu
Impacto Sobre o Risco de Crédito: Uma
Abordagem de Regressão Quantílica**

Kevin Amorim Gomes

30 de agosto de 2024

Kevin Amorim Gomes

Análise de Estresse Macroeconômico e seu Impacto Sobre o Risco de Crédito: Uma Abordagem de Regressão Quantílica

Trabalho de Conclusão apresentado à comissão de Graduação do Departamento de Estatística da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do título de Bacharel em Estatística.

Orientador: Prof. Dr. Hudson da Silva
Torrent

Porto Alegre
30 de agosto de 2024

Kevin Amorim Gomes

Análise de Estresse Macroeconômico e seu Impacto Sobre o Risco de Crédito: Uma Abordagem de Regressão Quantílica

Este Trabalho foi julgado adequado para obtenção dos créditos da disciplina Trabalho de Conclusão de Curso em Estatística e aprovado em sua forma final pela Orientador e pela Banca Examinadora.

Orientador: _____
Prof. Dr. Hudson da Silva Torrent, UFRGS
Doutor em Economia pela Universidade Federal do Rio Grande do Sul, Porto Alegre, RS

Banca Examinadora:

Prof. Dr. Marcio Valk, UNICAMP
Doutor em Estatística pela Universidade Estadual de Campinas - SP

Porto Alegre
30 de agosto de 2024

“Os dados são apenas resumos de milhares de histórias – conte algumas dessas histórias para ajudar a tornar os dados significativos.” (Chip & Dan Heath)

Agradecimentos

Agradeço a minha família, e em especial, ao meu pai, Gabriel, e minha mãe, Cristina, que me deram o devido apoio para eu chegar neste momento de conclusão de curso, desde caronas para pegar o ônibus e ir para o Campus do Vale, ou me buscar todos os dias no meu último semestre, vindo de Canoas até Porto Alegre só para isso.

Agradeço a minha namorada Ana Julia que me ajudou a passar pelos momentos mais difíceis da escrita durante os últimos meses, me ajudando sempre que possível para que eu não desistisse.

Agradeço aos meus colegas de curso, em especial aos 5 que me acompanharam durante toda a jornada: Giordano, Leonardo Damiani, João Lucas, Antônio e Letícia. Esse grupo se ajudou sempre que possível em trabalhos complicados e deixou tudo mais leve ao passar do tempo, mesmo tendo que fazer o curso durante uma pandemia.

Agradeço ao meu orientador Hudson por topar em entrar nessa jornada que durou mais de um ano, em um assunto que não tínhamos muita experiência mas sempre que necessário conseguiu me ajudar da melhor forma possível.

Por fim, agradeço a cada um que me ajudou a completar mais esta etapa da minha vida, além da oportunidade que a UFRGS e o Instituto de Matemática e Estatística me proporcionaram. Creio que além de aprender muito sobre estatística, vou levar ensinamentos sobre a vida que vão ajudar e muito no futuro.

Resumo

Este trabalho investiga a relação entre cenários de estresse em variáveis macroeconômicas e a taxa de inadimplência no crédito, utilizando modelos de regressão quantílica. A metodologia é estruturada em várias etapas, começando pela extração e construção de um banco de dados que inclui as variáveis macroeconômicas utilizadas para a análise, a realização de testes de estacionariedade para garantir que as séries temporais utilizadas sejam adequadas para análise e o modelo ARIMA é aplicado para prever as variáveis macroeconômicas. A parte central da análise utiliza a regressão quantílica, que é uma abordagem estatística que permite examinar como diferentes níveis de estresse macroeconômico afetam a inadimplência em diferentes quantis da distribuição de risco de crédito. Os resultados obtidos indicam que a taxa de inadimplência é relativamente influenciada por variáveis macroeconômicas, confirmando a hipótese de que cenários econômicos desfavoráveis levam a um aumento na inadimplência. Os cenários de estresse construídos demonstram que, em condições adversas, a inadimplência tende a aumentar, evidenciando a vulnerabilidade dos tomadores de crédito em períodos de crise econômica. Além disso, a análise revela que a taxa de inadimplência é mais fortemente impactada quando há estresse macroeconômico nas variáveis Selic/Over e concessões de crédito em comparação com cenários de estresse nas variáveis PIB, IPCA e desemprego. As limitações da pesquisa são discutidas, incluindo a necessidade de considerar outros fatores que podem influenciar a inadimplência e a possibilidade de expandir a análise para incluir diferentes contextos econômicos. Sugestões para futuras pesquisas incluem a exploração de outros modelos estatísticos e a inclusão de variáveis adicionais que possam enriquecer a compreensão da relação entre estresse macroeconômico e risco de crédito.

Palavras-Chave: Análises de estresse, Cenários de estresse, Carteiras de crédito, Teste de estacionariedade, Séries Temporais, ARIMA, Regressão quantílica, Econometria.

Abstract

This work investigates the relationship between stress scenarios in macroeconomic variables and credit default rates using quantile regression models. The methodology is structured in several stages, beginning with the extraction and construction of a database that includes the macroeconomic variables used for the analysis, conducting stationarity tests to ensure that the time series data are suitable for analysis, and applying the ARIMA model to forecast the macroeconomic variables. The core of the analysis uses quantile regression, which is a statistical approach that allows for examining how different levels of macroeconomic stress affect default rates across different quantiles of the credit risk distribution. The results indicate that the default rate is relatively influenced by macroeconomic variables, confirming the hypothesis that unfavorable economic scenarios lead to an increase in defaults. The constructed stress scenarios demonstrate that, under adverse conditions, defaults tend to increase, highlighting the vulnerability of borrowers during economic crises. Additionally, the analysis reveals that the default rate is more strongly impacted when there is macroeconomic stress in the Selic/Over rates and credit concessions variables, compared to stress scenarios in GDP, IPCA, and unemployment variables. The limitations of the research are discussed, including the need to consider other factors that may influence default rates and the possibility of expanding the analysis to include different economic contexts. Suggestions for future research include exploring other statistical models and incorporating additional variables that could enhance the understanding of the relationship between macroeconomic stress and credit risk.

Keywords: Stress Analysis, Stress Scenarios, Credit Portfolios, Stationarity Test, Time Series, ARIMA, Quantile Regression, Econometrics.

Sumário

1	Introdução	11
2	Revisão Bibliográfica	13
3	Metodologia	15
3.1	Teste de estacionariedade	15
3.2	Modelo ARIMA	16
3.3	Regressão Quantílica	16
3.4	Construção dos cenários	17
4	Banco de dados	20
4.1	Construção e manipulação	20
4.2	Previsão	24
5	Resultados	25
6	Conclusão	30
	Referências Bibliográficas	31

Lista de Figuras

Figura 4.1: Variáveis macroeconômicas iniciais em função do tempo	21
Figura 4.2: Variáveis macroeconômicas diferenciadas em função do tempo . . .	23
Figura 5.1: Densidade estimada por modelo - cenários multivariados	25
Figura 5.2: Densidade estimada por modelo - cenários para Selic/Over	26
Figura 5.3: Densidade estimada por modelo - cenários para PIB	26
Figura 5.4: Densidade estimada por modelo - cenários para IPCA	26
Figura 5.5: Densidade estimada por modelo - cenários para desemprego	27
Figura 5.6: Densidade estimada por modelo - cenários para concessões	27
Figura 5.7: ECDF estimada por cenários - Modelo 1	28
Figura 5.8: ECDF estimada por cenários - Modelo 2	29

Lista de Tabelas

Tabela 4.1: Testes ADF para as variáveis macroeconômicas utilizadas	22
Tabela 4.2: Valores críticos para o teste ADF	22
Tabela 4.3: Testes ADF para as variáveis macroeconômicas diferenciadas utilizadas	23
Tabela 4.4: Previsão para $H = 1$ das variáveis macroeconômicas independentes utilizadas	24

1 Introdução

Recentemente, houve um crescimento nos materiais sobre teste de estresse, definido por [Stein \(2012\)](#) como técnicas que tentam medir a sensibilidade da situação de uma carteira, instituição ou sistema financeiro contra variações de determinados fatores de risco. A importância de se encontrar uma metodologia adequada para esse problema foi destacada, recentemente, pela ocorrência de fortes crises, como a pandemia causada pela COVID-19, e por cascatas de falências de bancos em alguns países. As instituições financeiras atualmente utilizam esta ferramenta denominada teste de estresse para a gestão de riscos, uma vez que é de extrema importância para a avaliação do risco sistêmico.

De acordo com [Vazquez et al. \(2012\)](#), esses testes são realizados pelos bancos centrais e individuais para determinar potenciais fontes de risco que possam ser encontradas em cenários de mudança severa em condições macroeconômicas e testar sua capacidade de resistência nesses eventos. Este trabalho tem como tema a construção de cenários estressantes para variáveis macroeconômicas afim de uma análise onde os cenários construídos afetam a taxa de inadimplência de crédito geral, buscando analisar a possível hipótese na qual a taxa de inadimplência tem uma forte relação com o crescimento de variáveis macroeconômicas, e então é esperado que quanto pior o cenário, maior a taxa de inadimplência.

A abordagem utilizada neste trabalho baseia-se em [Schechtman e Gaglianone \(2012\)](#) e possui três passos principais. No primeiro passo, um modelo de regressão quantílica é estimado, relacionando inadimplência e variáveis macroeconômicas relevantes. Esse método permite entender o impacto do estresse macroeconômico sobre a distribuição da variável inadimplência, em particular na cauda da distribuição. No segundo passo, cada uma das variáveis macroeconômicas é modelada e obtida sua previsão um passo a frente. Finalmente, essas previsões e os coeficientes estimados no modelo de regressão quantílica são combinados, permitindo a construção de cenários de estresse macroeconômico sobre a taxa de inadimplência. Com essa metodologia, é possível avaliar como diferentes cenários macroeconômicos afetariam o risco de crédito dos bancos.

Além desta introdução, no capítulo seguinte é apresentado o referencial bibliográfico, contendo os modelos normalmente utilizados na literatura, bem como sua fundamentação teórica. O terceiro capítulo aborda toda a metodologia utilizada no trabalho, desde a estimação dos modelos de regressão quantílica até a construção dos cenários de estresse, obtidos com o auxílio dos modelos de previsão utilizados. O quarto capítulo aborda o banco de dados, mostrando a escolha das variáveis, as análises e manipulações realizadas, assim como a previsão calculada. Por fim, temos

o capítulo de resultados, que nos mostra os resultados e a interpretação final do modelo, permitindo o entendimento do impacto do estresse macroeconômico sobre a taxa de inadimplência. Por fim, é apresentado o capítulo de conclusão do trabalho, no qual apresentamos um resumo das principais implicações do trabalho, assim como suas limitações e possíveis abordagens futuras.

2 Revisão Bibliográfica

De acordo com [Vazquez et al. \(2012\)](#), os testes de estresse para o risco de crédito envolvem três tarefas principais:

1. O desenvolvimento de um modelo para capturar as inter-relações entre variáveis macroeconômicas e financeiras selecionadas. As variáveis macroeconômicas usadas habitualmente em modelos de teste de estresse incluem medidas de atividade econômica (ou seja, crescimento do PIB, hiato do produto e desemprego) e medidas de condições monetárias e preços-chave (ou seja, taxa de juros, taxa de câmbio, inflação, crescimento monetário e preços de imóveis);
2. A calibração de vetores de parâmetros que ligam variáveis macroeconômicas e financeiras a medidas específicas de desempenho de empréstimos;
3. O desenho de cenários macroeconômicos adversos e o cálculo de seu impacto na qualidade do crédito e na solvência dos bancos.

O objetivo desses testes é tornar os riscos mais transparentes, avaliando as perdas potenciais de uma determinada carteira em mercados anormais. Essas ferramentas são comumente usadas por instituições financeiras como parte de seus modelos internos e sistemas de gestão e para informar decisões sobre tomada de risco e alocação de capital. Além disso, essas ferramentas têm se tornado cada vez mais utilizadas pelos reguladores financeiros para avaliar a solidez dos sistemas financeiros sob seu controle, com vários artigos aplicando ferramentas de teste de estresse para avaliar a resiliência de vários sistemas bancários a choques macroeconômicos adversos. ([Barhill et al., 2006](#)) ([Busch et al., 2018](#)) ([Uquillas e Tonato, 2022](#)) ([Patra e Padhi, 2022](#))

Na ausência de modelos teóricos bem estabelecidos para explicar a relação entre as variáveis macroeconômicas e o risco de crédito, uma abordagem de forma reduzida e amplamente empregada na literatura aplicada é [Wilson \(1998\)](#). Nessa literatura, o objetivo principal é aferir a vulnerabilidade de uma carteira (mercado, crédito ou ambos) a cenários macroeconômicos adversos, ou a eventos ou choques extremos, mas plausíveis. Segundo [Schechtman e Gaglianone \(2012\)](#), a possível popularidade do modelo Wilson como portfólio de risco de crédito é atribuído ao fato da característica interna das variáveis macroeconômicas afetarem o risco de crédito. Por outro lado, [Schechtman e Gaglianone \(2012\)](#) também consideram um modelo alternativo para o risco de crédito que incorpora uma tendência estocástica do indicador de risco de crédito. Com isso, são realizados dois modelos e a comparação entre eles: o modelo tradicional de Wilson, amplamente utilizado na literatura para

analisar o risco de crédito, e um modelo de regressão quantílica, baseado em [Koenker e Xiao \(2002\)](#), utilizando-se das caudas de distribuição para representar o que de pior ainda pode acontecer ao risco de crédito condicionado a um cenário de variáveis macroeconômicas prejudicial.

Com isso, a abordagem a ser seguida neste trabalho foi o modelo de regressão quantílica, baseado em [Schechtman e Gaglianone \(2012\)](#). Essa abordagem possui uma característica atraente para análises de estresse: a importância relativa das variáveis macroeconômicas muda de acordo com o quantil da distribuição de risco de crédito. Em particular, variáveis macroeconômicas que têm um pequeno efeito relativo na mediana da distribuição podem ganhar relevância na explicação de um quantil alto do indicador de risco de crédito. Como modelo semiparamétrico, a QR relaxa uma suposição central de normalidade usada no modelo de Wilson, sendo a não-normalidade mais realista para exercícios de teste de estresse.

3 Metodologia

Este capítulo visa a explicação, passo a passo, de como o trabalho foi realizado, começando pela extração do banco de dados e tendo um fim na construção dos gráficos, que serão analisados posteriormente nos resultados. Para a realização deste trabalho, foi utilizada a linguagem de programação R através do software RStudio como ambiente de desenvolvimento.

O trabalho resumiu-se em cinco etapas, tais como a construção e manipulação do banco de dados, os testes de verificação da estacionariedade das séries utilizadas, o desenvolvimento dos modelos de previsão para as variáveis macroeconômicas, o desenvolvimento dos modelos de regressão quantílica e a construção dos gráficos através dos cenários utilizados para previsão da inadimplência.

3.1 Teste de estacionariedade

O teste de estacionariedade utilizado para verificar a presença de raiz unitária nas séries temporais das variáveis macroeconômicas foi o teste de Dickey-Fuller aumentado (ADF), proposto por [Dickey e Fuller \(1979\)](#). O teste ADF pode incluir um termo constante e um termo de tendência linear, ou ambos, de acordo com as características da série temporal. A equação relevante para o teste é dada por:

$$\Delta y_t = \mu + \delta_t + \alpha \cdot y_{t-1} + e_t,$$

Enquanto a hipótese nula e a hipótese alternativa do teste são:

$$H_0 : \alpha = 0 \text{ (presença de raiz unitária)}$$

$$H_1 : \alpha < 0 \text{ (série estacionária)}$$

Ou seja, o teste ADF com constante e tendência consiste em:

1. Estimar por MQO:

$$\Delta y_t = \mu + \delta_t + \alpha \cdot y_{t-1} + \beta_1 \cdot \Delta y_{t-1} + \beta_2 \cdot \Delta y_{t-2} + \dots + \beta_p \cdot \Delta y_{t-p} + e_t$$

2. Calcular a estatística do teste:

$$\text{Est. teste} = \frac{\hat{\alpha}}{\hat{\text{dp}}(\hat{\alpha})}$$

onde $\hat{dp}(\hat{\alpha})$ corresponde ao desvio padrão estimado para $\hat{\alpha}$.

Dessa forma, caso a estatística de teste for menor do que τ_α , rejeitamos H_0 , onde τ_α é o percentil de nível α da distribuição de Dickey-Fuller.

3.2 Modelo ARIMA

Para a realização da previsão das variáveis macroeconômicas utilizadas, o modelo autorregressivo integrado de médias móveis, mais conhecido como ARIMA, foi utilizado devido à sua ampla utilização na literatura e por ser uma metodologia bastante conhecida na modelagem de séries temporais. Segundo [Box George et al. \(1976\)](#), modelos para predição de séries temporais econômicas devem ser escolhidos com cuidado, e os modelos ARIMA demonstraram ser particularmente úteis numa ampla gama de situações práticas.

O modelo ARIMA é composto por três componentes principais: a parte autorregressiva do modelo, mais conhecida como AR, onde temos a variável regredida em seus próprios valores defasados; a parte de integração do modelo (I), ao qual significa a quantidade necessária de diferenciações para tornar a série estacionária; e a parte de médias móveis do modelo, também conhecida como MA, onde a variável é modelada em função dos erros defasados ([Matos, 1997](#)). Então, é possível utilizarmos a notação ARIMA(p, d, q), onde p é a ordem da parte AR, d é a ordem da parte de integração e q é a ordem da parte MA do modelo.

De acordo com [Morettin e Tolo \(2004\)](#), a equação geral do modelo ARIMA pode ser descrita por:

$$\left(1 - \sum_{i=1}^p \phi_i \cdot L^i\right)(1 - L)^d Y_t = \mu + \left(1 - \sum_{j=1}^q \theta_j \cdot L^j\right) e_t$$

onde:

- Y_t é o valor da série temporal no tempo t ;
- ϕ_i são os coeficientes da parte AR do modelo para cada defasagem i ;
- L é o operador de defasagem;
- d é o grau de diferenciação da série;
- μ é a constante;
- θ_j são os coeficientes da parte MA do modelo para cada defasagem j ;
- e_t é o termo de erro no tempo t .

3.3 Regressão Quantílica

Os modelos de regressão são extremamente úteis em estudos estatísticos pela sua facilidade de interpretação, normalmente utilizando-se do método de mínimos quadrados para estimar a média condicional de uma variável de interesse. Porém, o

interesse em alguns casos pode ser outro, como a mediana ou, neste caso, os quantis da distribuição condicional.

Com isso, o modelo utilizado foi o de regressão quantílica, desenvolvido por [Koenker e Bassett Jr \(1978\)](#) e que permite estimar os efeitos das variáveis independentes em diferentes quantis da distribuição condicional da variável dependente, oferecendo uma análise mais abrangente do que a regressão linear tradicional.

A regressão quantílica pode ser formalmente expressa pela seguinte equação:

$$Q_\tau(y_i|x_{i1}, x_{i2}, \dots, x_{ip}) = \beta_0(\tau) + x_{i1}\beta_1(\tau) + x_{i2}\beta_2(\tau) + \dots + x_{ip}\beta_p(\tau)$$

onde:

- $Q_\tau(y_i|x_{1,i}, x_{2,i}, \dots, x_{p,i})$ é o τ -ésimo quantil da variável dependente condicional as variáveis independentes;
- τ representa o quantil de interesse, com $0 < \tau < 1$;
- $\beta_k(\tau)$ é o coeficiente k a ser estimado para o quantil τ .

Para estimar $\beta_k(\tau)$, minimizamos a seguinte função de perda quantílica:

$$\min_{\beta_0(\tau), \dots, \beta_p(\tau)} \sum_{i=1}^n \rho_\tau \left(y_i - \beta_0(\tau) - \sum_{j=1}^p x_{ij}\beta_j(\tau) \right)$$

onde:

- $\rho_\tau(u)$ é a função de perda quantílica definida como $\rho_\tau(u) = u(\tau - I(u < 0))$;
- $I(\cdot)$ é a função indicadora que vale 1 se $u < 0$ e 0 caso contrário.

3.4 Construção dos cenários

O primeiro passo após a construção do banco de dados foi realizar o teste de estacionariedade para as variáveis macroeconômicas utilizadas e explicadas no capítulo 4. Para todas as variáveis, foi utilizado o teste considerando apenas o intercepto, e o critério de informação utilizado para determinar o número de defasagens foi o critério Bayesiano de Informação (BIC). A função utilizada no R foi a `ur.df`, disponível na biblioteca `Urca`.

Com isso, foi possível realizar a estimação dos modelos de regressão quantílica. Os modelos estimados para nossa variável dependente (inadimplência da carteira de crédito geral) foram baseados em [Schechtman e Gaglianone \(2012\)](#). Foram realizados dois modelos de regressão quantílica, onde o primeiro modelo tem como variáveis independentes as variáveis macroeconômicas e o lag de primeiro passo da inadimplência, e o segundo modelo têm as variáveis macroeconômicas, o lag de primeiro passo das variáveis macroeconômicas e o lag de primeiro passo da inadimplência.

Modelo 1:

$$\begin{aligned} Q_\tau(y_t|y_{t-1}, \text{SelicOver}_t, \text{PIB}_t, \text{IPCA}_t, \text{desemprego}_t, \text{concessões}_t) = & \beta_0(\tau) + y_{t-1}\beta_1(\tau) \\ & + \text{SelicOver}_t\beta_2(\tau) + \text{PIB}_t\beta_3(\tau) + \text{IPCA}_t\beta_4(\tau) + \text{desemprego}_t\beta_5(\tau) \\ & + \text{concessões}_t\beta_6(\tau); \tau \in [0, 1] \end{aligned}$$

Modelo 2:

$$\begin{aligned}
Q_\tau(y_t|y_{t-1}, \text{SelicOver}_t, \text{PIB}_t, \text{IPCA}_t, \text{desemprego}_t, \text{concessões}_t, \\
\text{SelicOver}_{t-1}, \text{PIB}_{t-1}, \text{IPCA}_{t-1}, \text{desemprego}_{t-1}, \text{concessões}_{t-1}) = \beta_0(\tau) + y_{t-1}\beta_1(\tau) \\
+ \text{SelicOver}_t\beta_2(\tau) + \text{PIB}_t\beta_3(\tau) + \text{IPCA}_t\beta_4(\tau) + \text{desemprego}_t\beta_5(\tau) \\
+ \text{concessões}_t\beta_6(\tau) + \text{SelicOver}_{t-1}\beta_7(\tau) + \text{PIB}_{t-1}\beta_8(\tau) + \text{IPCA}_{t-1}\beta_9(\tau) \\
+ \text{desemprego}_{t-1}\beta_{10}(\tau) + \text{concessões}_{t-1}\beta_{11}(\tau); \tau \in [0, 1]
\end{aligned}$$

Para a construção dos modelo ARIMA para cada variável, foi utilizada a função `auto.arima` disponível na biblioteca `forecast`. A função utiliza-se do método desenvolvido por [Hyndman e Khandakar \(2008\)](#), que combina testes de raiz unitária, minimização dos critérios de informação realizados pela função (AIC, BIC e AICc) e o método de máxima verossimilhança para obter o melhor modelo ARIMA. Vale ressaltar também que, devido a realização dos testes de estacionariedade, os modelos construídos e utilizados foram apenas modelos ARMA, pois as variáveis utilizadas não necessitaram de diferenciação.

Então, após os testes de estacionariedade utilizados para verificarmos a raiz unitária das variáveis, o método de construção do modelo ARMA que nos permite prever os valores das variáveis macroeconômicas um passo a frente e os modelos de regressão quantílica utilizados para estimar a distribuição condicional da variável dependente com diferentes quantis (nesse caso, entre 0 e 1), para realizarmos uma análise de estresse na variável de inadimplência de crédito geral, foi necessário executar a construção de cenários de estresse para as variáveis macroeconômicas utilizadas no modelo de regressão quantílica.

Em [Schechtman e Gaglianone \(2012\)](#), é realizada uma construção de cenários em um horizonte de tempo de $H = 1$ até $H = 4$ para as variáveis macroeconômicas de interesse. [Schechtman e Gaglianone \(2012\)](#) também realizam um modelo VAR para as variáveis macroeconômicas e utilizam-se das últimas observações disponíveis, assim como a soma ou subtração de 1, 2 ou 3 desvios padrões das estimações para construção dos cenários. Para nosso trabalho, a construção dos cenários é para apenas um passo frente ($H = 1$), e utiliza-se também de 1, 2 ou 3 desvios padrões, calculados na estimação dos modelos ARMA já citados e que serão detalhados posteriormente na seção 4.2. Com isso, os cenários construídos foram:

- **Cenário sem estresse:** previsão obtida através do modelo ARMA é utilizada na regressão quantílica sem a soma ou subtração de desvio padrão;
- **Cenário 1:** previsão obtida através do modelo ARMA é utilizada na regressão quantílica com a soma ou subtração de 1 desvio padrão;
- **Cenário 2:** previsão obtida através do modelo ARMA é utilizada na regressão quantílica sem a soma ou subtração de 2 desvios padrões;
- **Cenário 3:** previsão obtida através do modelo ARMA é utilizada na regressão quantílica sem a soma ou subtração de 3 desvios padrões;

Então, para o cálculo dos cenários, foi realizada uma simulação com 2000 valores da regressão quantílica para os dois modelos com $\tau \sim U(0, 1)$ com o valor de previsão

de $H = 1$ sem a adição ou subtração de desvio padrão, ou adicionado ou subtraído por um, dois ou três desvios padrões, dado o coeficiente de cada variável, para analisarmos o efeito das variáveis estressadas na variável dependente. Quando o coeficiente da variável no modelo de regressão era negativo, o desvio padrão era subtraído. Quando o coeficiente da variável no modelo era positivo, o desvio padrão era adicionado.

Os modelos com cenários construídos foram:

- **Modelo multivariado:** adição ou subtração de um, dois ou três desvios padrões de $H = 1$ para todas as variáveis macroeconômicas do modelo;
- **Modelo com cenários para Selic/Over:** adição ou subtração de um, dois ou três desvios padrões para a previsão de $H = 1$ da taxa de crescimento Selic/Over mensal;
- **Modelo com cenários para o PIB:** adição ou subtração de um, dois ou três desvios padrões para a previsão de $H = 1$ da taxa de crescimento do PIB mensal;
- **Modelo com cenários para IPCA:** adição ou subtração de um, dois ou três desvios padrões para a previsão de $H = 1$ da taxa de crescimento do IPCA mensal;
- **Modelo com cenários para desemprego:** adição ou subtração de um, dois ou três desvios padrões para a previsão de $H = 1$ da taxa de crescimento de desemprego mensal;
- **Modelo com cenários para concessões:** adição ou subtração de um, dois ou três desvios padrões para a previsão de $H = 1$ da taxa de crescimento de concessões mensal;

Com isso, a previsão dos modelos ARMA aplicada a regressão quantílica no exercício de simulação permite a construção da distribuição da variável dependente condicional às variáveis independentes, utilizando diferentes cenários de estresse. Para analisarmos como os cenários de estresse afetam a taxa de inadimplência geral, foi calculada a função de distribuição empírica (ECDF) dos valores preditos para a inadimplência em cada um desses cenários. A ECDF será explicada posteriormente para melhor entendimento na seção 5.

4 Banco de dados

4.1 Construção e manipulação

Para a construção do banco de dados, foram utilizadas as variáveis macroeconômicas iniciais, providas de [IPEADATA \(2023\)](#): taxa de juros Over/Selic acumulada no mês, a taxa de variação do IPCA (Índice Nacional de Preços ao Consumidor amplo), a variação em % no período anterior do PIB (Produto Interno Bruto) brasileiro deflacionado pela taxa IGP-DI, a variação em % no período anterior de concessões de crédito também deflacionada pela taxa IGP-DI, e a inadimplência da carteira de crédito geral. A fonte para todas as variáveis é o Banco Central do Brasil (BACEN), exceto pela taxa de variação do IPCA, provida do Instituto Brasileiro de Geografia e Estatística (IBGE). Também foi utilizada a taxa de desocupação brasileira das pessoas de 14 anos ou mais, retiradas de estimativas realizadas em [Lameiras e Hecksher \(2023\)](#).

Estes dados foram extraídos de forma mensal, a partir de janeiro de 2012 até setembro de 2023, totalizando 141 linhas. A data inicial escolhida foi devido à disponibilidade dos dados, uma vez que a taxa de desemprego utilizada atualmente no país começou no mesmo período citado. Também é válido ressaltar algumas mudanças após a construção, como a utilização do logaritmo da taxa de juros Selic/Over e o logaritmo da taxa de desocupação. Também, conforme [Schechtman e Gaglianone \(2012\)](#), foi aplicada uma transformação logito na taxa de inadimplência da carteira de crédito geral para a realização dos modelos, uma vez que a transformação é comum para lidar com proporções e ajudando na linearização de relações, estabilização de variâncias e persistência da série.

Estas variáveis macroeconômicas foram escolhidas pois capturam diferentes aspectos da saúde econômica e financeira de um país: a taxa Selic/Over é a principal taxa de juros da economia brasileira e influencia diretamente o custo de crédito; o PIB é um indicador fundamental da atividade econômica; o IPCA mede a inflação de um país; a taxa de desemprego serve como um indicador direto da situação do mercado de trabalho; e a concessão de créditos está ligada à disponibilidade de financiamento na economia. ([Bernanke e Gertler, 1995](#)) ([Mankiw, 2014](#)) ([Friedman, 1963](#)) ([Nickell e Layard, 1999](#)) ([Rajan, 1994](#))

Com isso, seguiu-se a análise descritiva dos dados e possível manipulação devido a alguma série apresentar não estacionariedade. Abaixo, temos os gráficos das variáveis macroeconômicas em função do tempo:

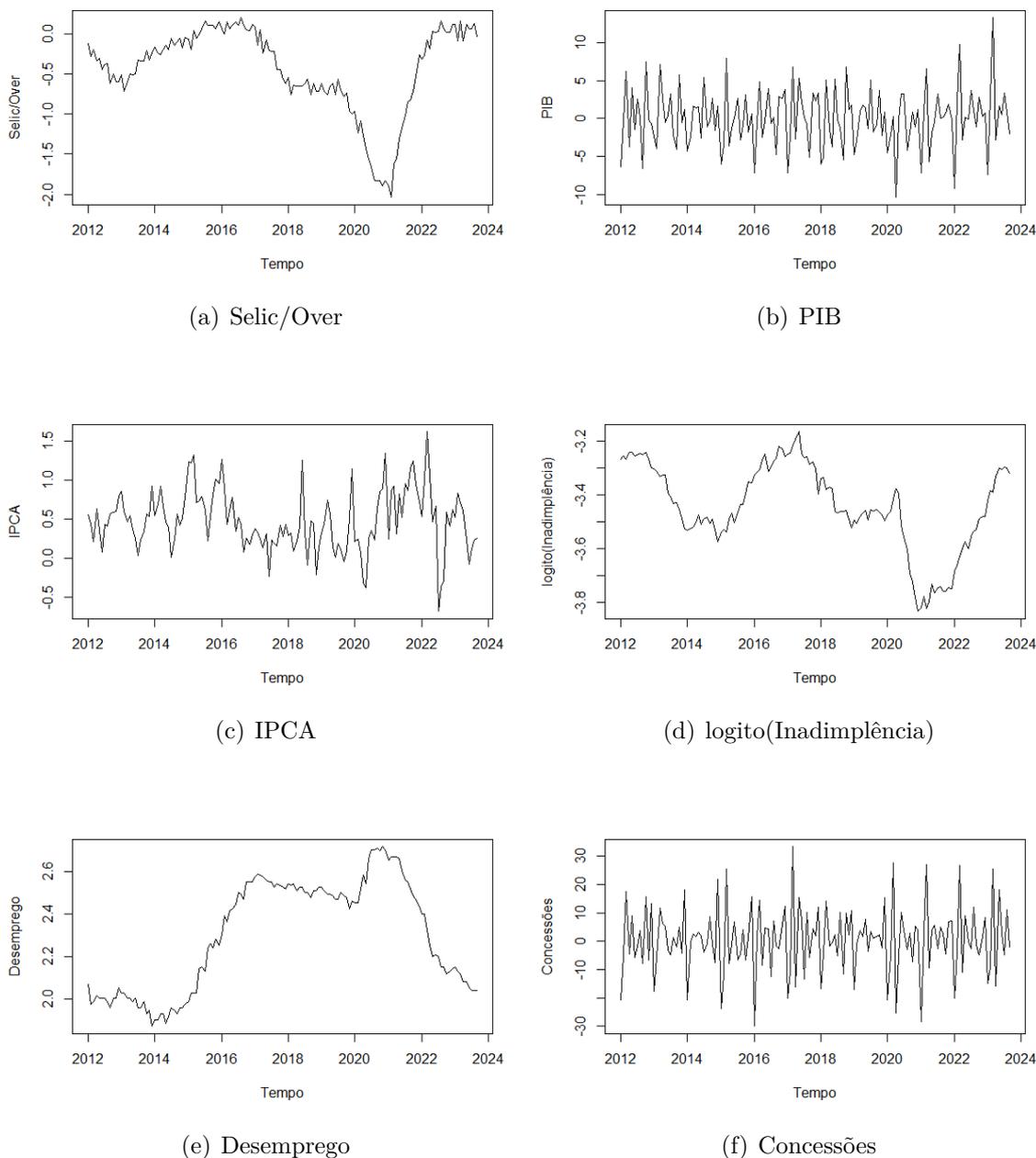


Figura 4.1: Variáveis macroeconômicas iniciais em função do tempo

Fonte: Autor.

Analisando pelos gráficos, temos que as séries do PIB, IPCA e concessões de crédito apresentam um comportamento estacionário, enquanto as séries temporais da taxa Selic/Over, desemprego e o logito da inadimplência apresentam um comportamento não estacionário. Para confirmarmos essas suposições, foram realizados os testes ADF descritos anteriormente em para cada uma das séries:

Tabela 4.1: Testes ADF para as variáveis macroeconômicas utilizadas

Variáveis	Estatística ADF
Selic/Over	-0.8497
PIB	-13.0966
IPCA	-5.1234
Desemprego	-0.9601
Concessões	-17.5614
Inadimplência	-1.5787

Fonte: Autor.

Tabela 4.2: Valores críticos para o teste ADF

p-valor	Valor
0.01	-3.46
0.05	-2.88
0.10	-2.57

Fonte: Autor.

Ao analisarmos estes valores, podemos ver que as variáveis Selic/Over, desemprego e inadimplência não rejeitam a hipótese nula de raiz unitária, ou seja, não são séries estacionárias. Para estas séries, foi necessário realizar a primeira diferença, mudando a interpretação das séries (taxa de crescimento mensal da taxa Selic/Over acumulada no mês, taxa de crescimento da desocupação mensal e taxa de crescimento da inadimplência geral). Com isso, foram realizados os gráficos das primeiras diferenças das séries temporais e o teste ADF para analisarmos a estacionariedade das séries:

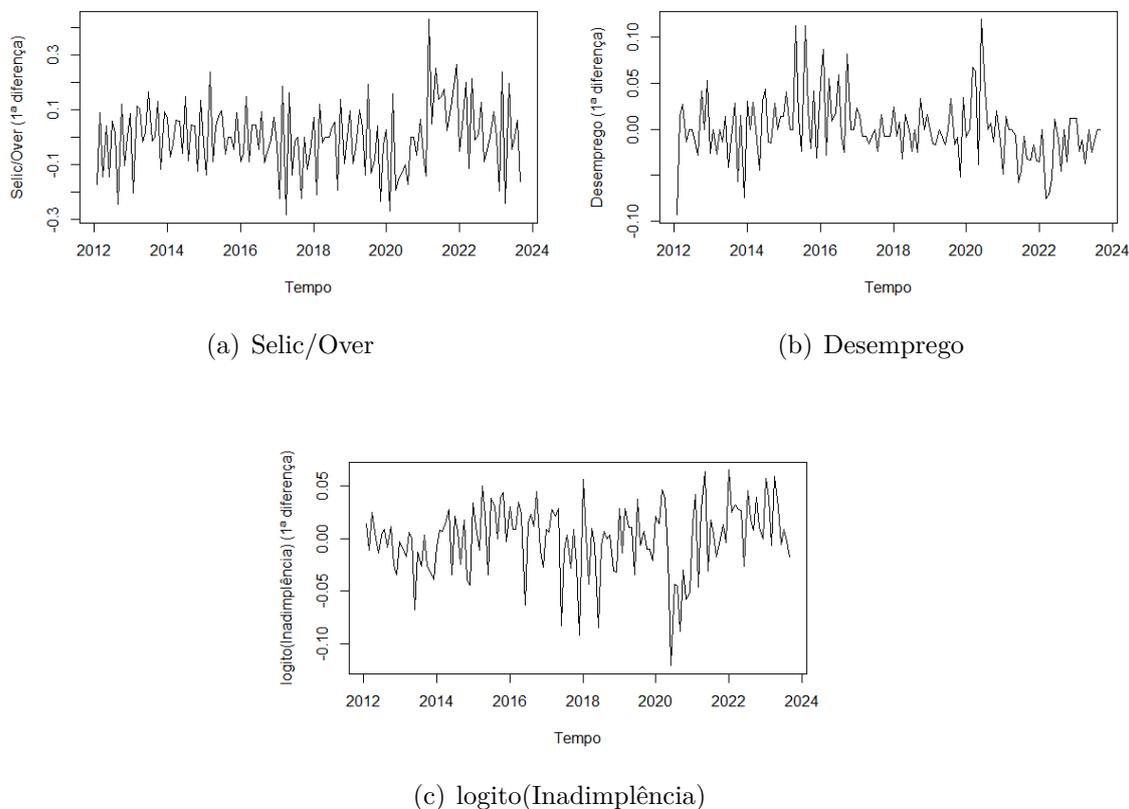


Figura 4.2: Variáveis macroeconômicas diferenciadas em função do tempo
Fonte: Autor.

Tabela 4.3: Testes ADF para as variáveis macroeconômicas diferenciadas utilizadas

Variáveis (1ª diferença)	Estatística ADF
Selic/Over	-7.8356
Desemprego	-7.3938
Inadimplência	-7.2799

Fonte: Autor.

Ao analisarmos os gráficos e a estatística de teste, rejeitamos a hipótese nula de raiz unitária da primeira diferença das séries analisadas. Então, o nosso banco de dados utilizado ao longo de todo o estudo segue como:

- **Selic/Over:** taxa de crescimento mensal acumulada no mês;
- **PIB:** taxa de crescimento mensal deflacionado pela taxa IGP-DI;
- **IPCA:** taxa de crescimento mensal;
- **Desemprego:** taxa de crescimento mensal (pessoas de 14 anos ou mais);
- **Concessões:** taxa de crescimento mensal deflacionado pela taxa IGP-DI;
- **Inadimplência:** taxa de crescimento do logito da carteira de crédito geral.

4.2 Previsão

Para o cálculo da previsão, foi necessário estimar os modelos ARMA para as variáveis macroeconômicas, através da função `auto.arima`, citada na seção 3.4. Com isso, os modelos finais para cada uma das variáveis utilizadas foram:

- **Selic/Over:** ARMA(3, 1) com média 0;
- **PIB:** ARMA(1, 2) com média 0;
- **IPCA:** AR(1) com média diferente de 0;
- **Desemprego:** ARMA(1, 2) com média 0;
- **Concessões:** ARMA(2, 2) com média diferente de 0.

Após a estimação dos modelos, foi realizado o cálculo da previsão para $H = 1$ dos modelos estimados. Para isso, foi utilizada a função `forecast`, da biblioteca `forecast`. Então, foi possível estimar os valores de previsão para $H = 1$ de cada uma das variáveis macroeconômicas, junto ao seu desvio padrão:

Tabela 4.4: Previsão para $H = 1$ das variáveis macroeconômicas independentes utilizadas

Variáveis	Média (desvio padrão)
Selic/Over	0.083 (0.10093)
PIB	0.656 (3.46505)
IPCA	0.351 (0.31705)
Desemprego	-0.0097 (0.031491)
Concessões	0.729 (6.41513)

Fonte: Autor.

5 Resultados

Os gráficos a seguir apresentam a estimativa da densidade da taxa de inadimplência referente a cada um dos dois modelos e a cada cenário de estresse macroeconômico considerado. Para isso, como dito anteriormente, foi necessário estimar os modelos de regressão quantílica através da simulação de 2000 valores com $\tau \sim U(0, 1)$.

Abaixo, temos a densidade com os cenários para todas as variáveis estressadas (modelo multivariado) em ambos os modelos de regressão quantílica:

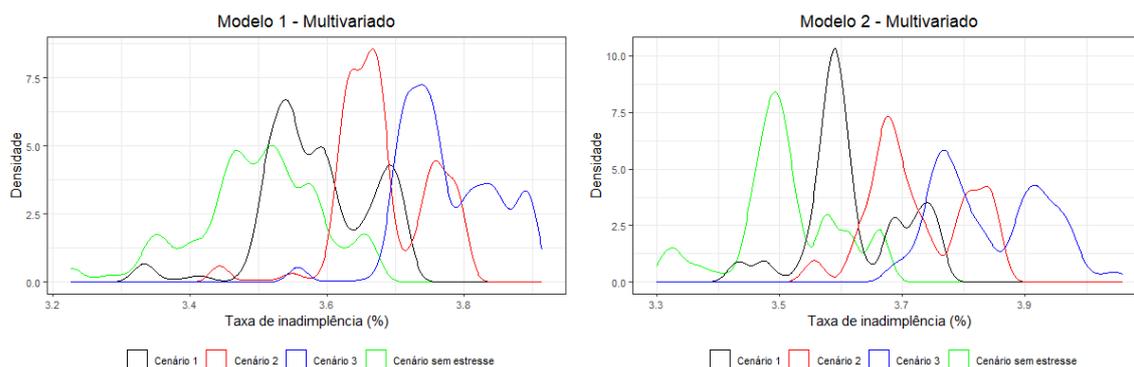


Figura 5.1: Densidade estimada por modelo - cenários multivariados
Fonte: Autor.

Analisando os resultados demonstrados, a figura 5.1 apresenta que para o modelo multivariado, para qualquer uma das regressões quantílicas, os valores de taxa de inadimplência da carteira de crédito geral tendem a uma maior densidade para valores maiores. Essa afirmação é fortificada pela média das curvas do modelo 1, por exemplo: 3.44 para o cenário sem estresse, 3.52 para o cenário 1, 3.62 para o cenário 2 e 3.74 para o cenário 3.

Então, a seguir, temos os gráficos para todos os modelos realizados com estresse em apenas uma variável, dividido por cenários, em ambos os modelos de regressão quantílica:

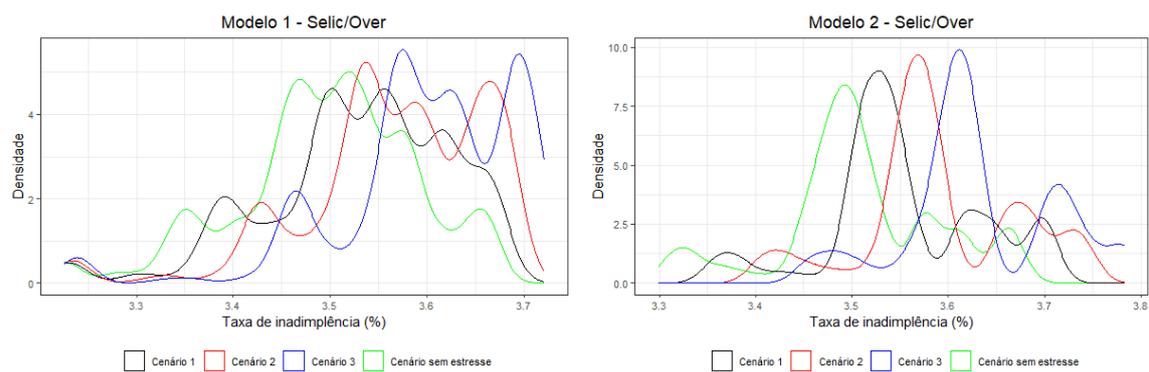


Figura 5.2: Densidade estimada por modelo - cenários para Selic/Over
Fonte: Autor.

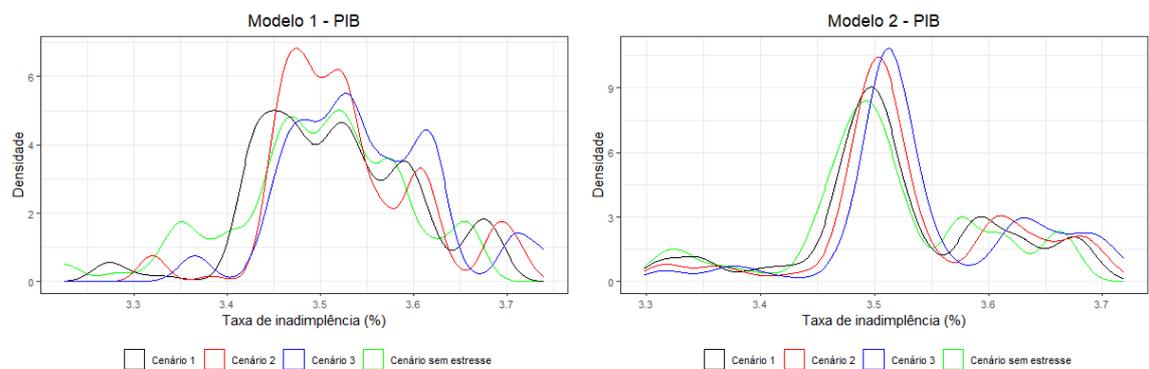


Figura 5.3: Densidade estimada por modelo - cenários para PIB
Fonte: Autor.

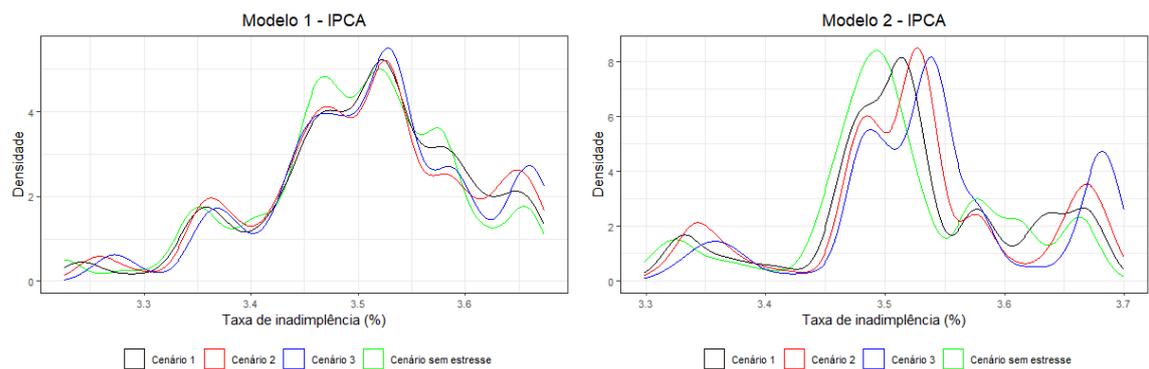


Figura 5.4: Densidade estimada por modelo - cenários para IPCA
Fonte: Autor.

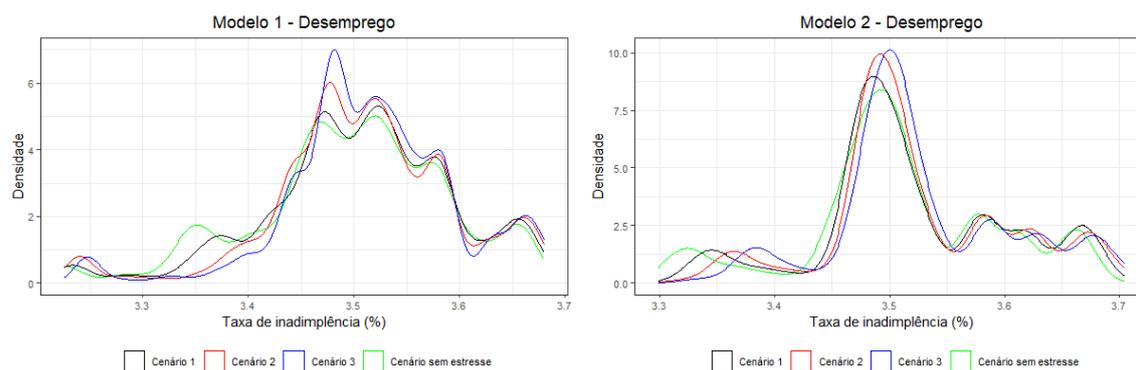


Figura 5.5: Densidade estimada por modelo - cenários para desemprego

Fonte: Autor.

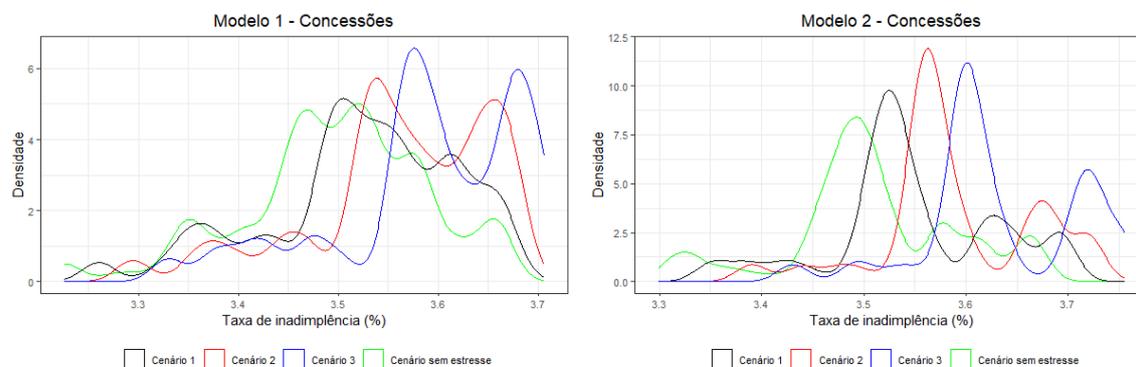


Figura 5.6: Densidade estimada por modelo - cenários para concessões

Fonte: Autor.

Analisando as densidades estimadas, podemos ver que as variáveis com maior diferença entre os diferentes cenários com estresse são a Selic/Over e a concessão de crédito, enquanto as outras variáveis não apresentam um comportamento tão diferente em relação a cada um dos cenários. É válido ressaltar que essa diferença é relativa, uma vez que estamos comparando a mesma variável dentro do modelo para possíveis cenários de estresse.

Após os gráficos de densidade para cada um dos modelos e cenários, foi estimada a curva ECDF referente a cada um dos cenários nos modelos de regressão quantílica realizados. Nesse caso, a curva nos ajuda a estimar qual a probabilidade em que uma taxa de inadimplência da carteira de crédito geral seja menor ou igual a um determinado valor, sem considerarmos uma distribuição paramétrica para os dados.

Em seguida, temos os gráficos para os três cenários realizados no modelo 1, divididos pelos cenários construídos para cada variável, para o modelo multivariado e para o cenário sem estresse:

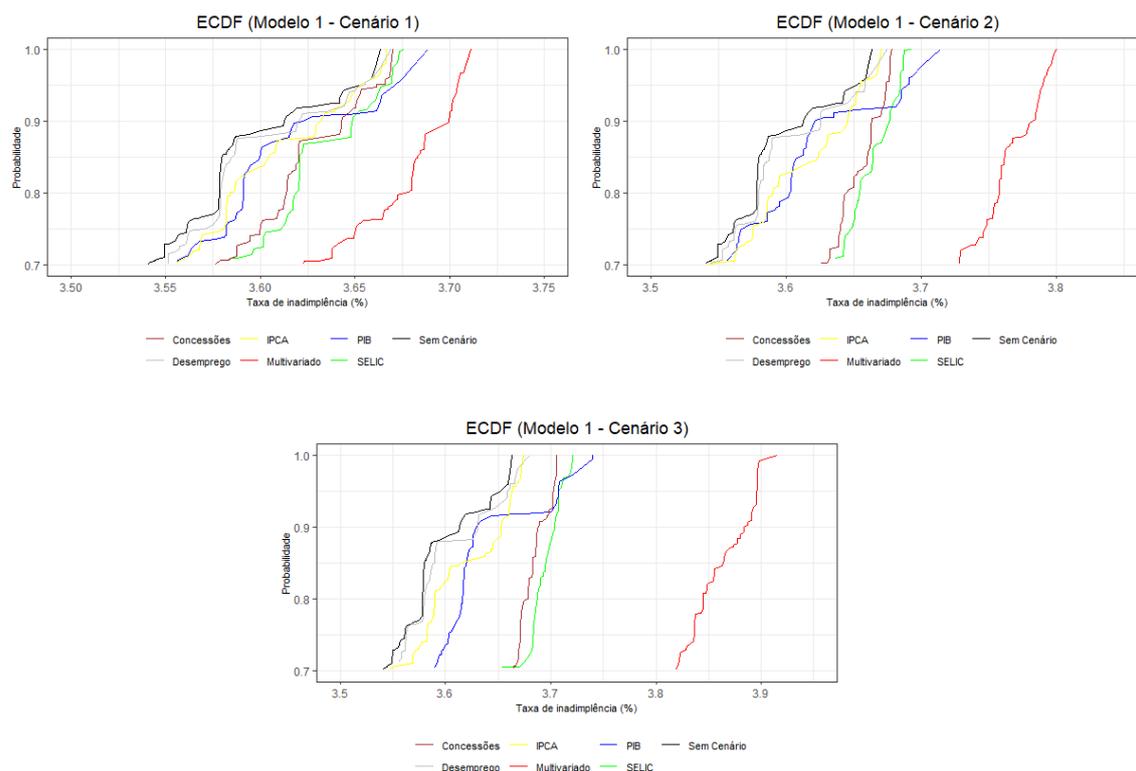


Figura 5.7: ECDF estimada por cenários - Modelo 1
Fonte: Autor.

Para os três cenários do modelo 1, podemos ver que sempre há um aumento do modelo multivariado em relação aos outros modelos. Analisando os valores, temos:

- Modelo multivariado: para o cenário 1, 95% das estimativas de taxa de inadimplência são menores ou iguais a 3.7. Para o cenário 2, 95% das estimativas de taxa de inadimplência são menores ou iguais a 3.78. E para o cenário 3, 95% das estimativas de taxa de inadimplência são menores ou iguais a 3.89. Essa tendência de crescimento dos valores nos indica um aumento da taxa de inadimplência em função dos cenários de estresse para as variáveis macroeconômicas.

É possível realizar também uma comparação no mesmo cenário entre diferentes variáveis macroeconômicas. Por exemplo, para o modelo 1 no cenário 1, o modelo de regressão quantílica sem estresse nos diz que em 95% das estimativas da taxa de inadimplência são menores ou iguais a 3.65, enquanto que o cenário 1 para o modelo da Selic/Over nos diz que em 95% das estimativas da taxa de inadimplência são menores ou iguais a 3.67, e como havíamos dito anteriormente, o modelo multivariado apresenta um valor máximo igual a 3.7 em 95% dos casos.

Analisando os gráficos, um possível padrão a se observar é que, relacionando as curvas ECDF com as densidades estimadas, identifica-se que o modelo multivariado sempre tem uma tendência de valores um pouco maior que os outros modelos. Também é possível notar que, dentre os modelos com variação em apenas uma variável macroeconômica comparadas ao modelo sem estresse, as variáveis Selic/Over e concessão de crédito se destacam como aquelas que mais impactam a inadimplência, gerando mais estresse econômico.

Em seguida, temos o segundo modelo de regressão quantílica realizado para cada um dos cenários, e estratificado pelas variáveis macroeconômicas:

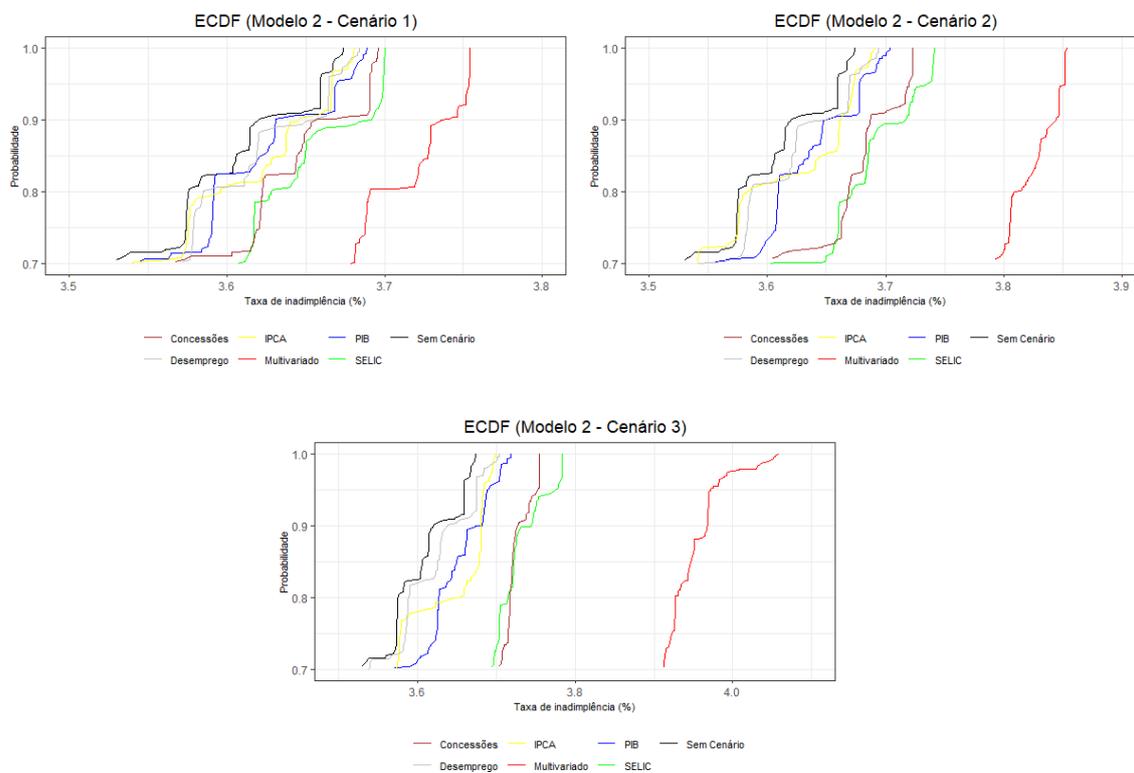


Figura 5.8: ECDF estimada por cenários - Modelo 2
Fonte: Autor.

Comparando os modelos, é possível analisarmos que, utilizando como exemplo o modelo 1 e multivariado para o cenário 3, em 95% dos casos, este apresenta uma taxa de inadimplência menor ou igual a 3.89, enquanto o modelo 2 e multivariado para o cenário 3 apresenta uma taxa menor ou igual a 3.98. Com isso, vemos que houve um aumento relativo da inadimplência de aproximadamente 2.31% do modelo 2 em relação ao modelo 1.

6 Conclusão

Este trabalho estimou dois modelos de regressão quantílica utilizando cenários de estresse através de previsões calculadas para cada uma das variáveis macroeconômicas. A primeira parte do trabalho utilizou de modelos ARMA para tratarmos com as variáveis macroeconômicas e explorar as previsões que estas variáveis podem proporcionar, dependendo apenas delas mesmas. A segunda parte utilizou-se de um modelo de regressão quantílica utilizando estas previsões e também os cenários construídos para analisarmos o efeito destas variáveis na taxa de inadimplência de crédito geral. Com os resultados, foi possível analisarmos que o modelo multivariado onde os cenários são utilizados em todas as variáveis apresenta maiores valores em relação aos outros modelos, seguido pelo modelo onde apenas o cenário da variável Selic/Over é utilizado e do modelo onde apenas o cenário da variável de concessões de crédito é utilizado. Isso corrobora parcialmente com o que [Schechtman e Gaglianone \(2012\)](#) apresenta, onde os modelos multivariados tem valores maiores, porém o cenário para algumas variáveis como o desemprego e o PIB são piores em relação a taxa de inadimplência com os resultados obtidos.

É possível notar também que, considerando uma relação entre os modelos, o modelo 2, que considera tanto a previsão estimada como o primeiro passo para trás das variáveis macroeconômicas no modelo de regressão quantílica, apresentou maiores valores de taxa de inadimplência comparados ao modelo 1, que utiliza apenas as previsões das variáveis macroeconômicas. Uma possível hipótese para estes resultados é de que a incorporação de informações adicionais sobre o passado das variáveis macroeconômicas pode melhorar a capacidade preditiva do modelo de regressão.

Uma das possíveis limitações para este trabalho é que não é possível afirmarmos que existe uma diferença significativa entre os modelos, uma vez que não há nenhum teste estatístico realizado para testarmos a diferença dos resultados. Então, o trabalho aponta apenas possíveis interpretações dos resultados, dado que temos uma diferença entre os valores estimados pelos modelos, mas não é possível calcular a significância desta diferença. Esta limitação se deve ao fato de que nessa literatura, não é comum a utilização de testes estatísticos para apresentar se há diferença significativa entre os resultados.

Outra possível limitação é de que nas principais literaturas, a predição calculada para as variáveis macroeconômicas utiliza-se de um modelo vetorial autorregressivo (VAR) ao invés do modelo ARMA utilizado para cada uma das variáveis. O modelo VAR poderia fornecer uma abordagem mais robusta e abrangente para analisar e prever variáveis macroeconômicas, pois consegue capturar interações entre as variáveis e estudar a causalidade e o impacto de choques.

Uma possível contribuição prática deste trabalho é considerar a influência de cenários macroeconômicos na taxa de inadimplência geral de crédito em um país. Isto ajuda na projeção econômica nacional, onde as variáveis que definem a situação econômica influenciam completamente na situação de crédito.

Referências Bibliográficas

- Barnhill, T. M., Souto, M. R., Tabak, B. M., et al. (2006). An analysis of off-site supervision of banks' profitability, risk and capital adequacy: a portfolio simulation approach applied to brazilian banks. Technical report.
- Bernanke, B. S. e Gertler, M. (1995). Inside the black box: the credit channel of monetary policy transmission. *Journal of Economic perspectives*, 9(4):27–48.
- Box George, E., Jenkins Gwilym, M., Reinsel Gregory, C., e Ljung Greta, M. (1976). Time series analysis: forecasting and control. *San Francisco: Holden Bay*.
- Busch, R., Koziol, P., e Mitrovic, M. (2018). Many a little makes a mickle: Stress testing small and medium-sized german banks. *The Quarterly Review of Economics and Finance*, 68:237–253.
- Dickey, D. A. e Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.
- Friedman, M. (1963). Inflation: causes and consequences. first lecture. *Repr. in Dollars and Deficits: Inflation, Monetary Policy and the Balance of Payments*, pages 21–46.
- Hyndman, R. J. e Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for r. *Journal of statistical software*, 27:1–22.
- IPEADATA (2023). Ipeadata. <http://www.ipeadata.gov.br>. Acesso em: novembro de 2023.
- Koenker, R. e Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50.
- Koenker, R. e Xiao, Z. (2002). Inference on the quantile regression process. *Econometrica*, 70(4):1583–1612.
- Lameiras, M. A. P. e Hecksher, M. (2023). Indicadores mensais do mercado de trabalho - setembro de 2023. <https://www.ipea.gov.br/cartadeconjuntura/index.php/2023/11/indicadores-mensais-do-mercado-de-trabalho-setembro-de-2023/>. Acesso em: novembro de 2023.
- Mankiw, N. G. (2014). *Macroeconomics*. Worth Publishers.

- Matos, O. C. (1997). *Econometria básica: teoria e aplicações*. Atlas.
- Morettin, P. A. e Tolo, C. M. d. C. (2004). *Análise de séries temporais*. Edgard Blucher.
- Nickell, S. e Layard, R. (1999). Labor market institutions and economic performance. *Handbook of labor economics*, 3:3029–3084.
- Patra, B. e Padhi, P. (2022). Resilience of indian banks: Macroeconomic stress test modeling for credit risk. *Journal of Public Affairs*, 22(1):e2350.
- Rajan, R. G. (1994). Why bank credit policies fluctuate: A theory and some evidence. *the Quarterly Journal of economics*, 109(2):399–441.
- Schechtman, R. e Gaglianone, W. P. (2012). Macro stress testing of credit risk focused on the tails. *Journal of Financial Stability*, 8(3):174–192.
- Stein, R. M. (2012). The role of stress testing in credit risk management. *Journal of investment management*, 10(4):64.
- Uquillas, A. e Tonato, R. (2022). Inter-portfolio credit risk contagion including macroeconomic and financial factors: A case study for ecuador. *Economic Analysis and Policy*, 73:299–320.
- Vazquez, F., Tabak, B. M., e Souto, M. (2012). A macro stress test model of credit risk for the brazilian banking sector. *Journal of Financial Stability*, 8(2):69–83.
- Wilson, T. C. (1998). Portfolio credit risk. *Economic policy review*, 4(3).