

UNIVERSIDADE FEDERAL DO CEARÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA - CAEN
MESTRADO PROFISSIONAL EM ECONOMIA DO SETOR PÚBLICO

FELIPE LEVI OLIVEIRA NOGUEIRA

PREVISÃO DO ICMS DO CEARÁ A PARTIR DOS CNAES DE ARRECADAÇÃO
DOS ANOS DE 2011 A 2022.

FORTALEZA

2024

FELIPE LEVI OLIVEIRA NOGUEIRA

PREVISÃO DO ICMS DO CEARÁ A PARTIR DOS CNAES DE ARRECADAÇÃO
DOS ANOS DE 2011 A 2022.

Dissertação apresentada ao Mestrado Profissional em Economia do Setor Público do Departamento de Economia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em economia.

Orientador: Prof. Dr. Roberto Tatiwa Ferreira

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

N712p Nogueira, Felipe Levi Oliveira.
Previsão do ICMS do Ceará a partir dos CNAEs de arrecadação dos anos de 2011 a 2022. / Felipe Levi Oliveira Nogueira. – 2024.
33 f.

Dissertação (mestrado) – Universidade Federal do Ceará, Faculdade de Economia, Administração, Atuária e Contabilidade, Mestrado Profissional em Economia do Setor Público, Fortaleza, 2024.
Orientação: Prof. Dr. Roberto Tatiwa Ferreira.

1. FAVAR. 2. ICMS. 3. Previsão de receita. 4. CNAE. I. Título.

CDD 330

FELIPE LEVI OLIVEIRA NOGUEIRA

PREVISÃO DO ICMS DO CEARÁ A PARTIR DOS CNAES DE ARRECADAÇÃO
DOS ANOS DE 2011 A 2022.

Dissertação apresentada ao Mestrado Profissional em Economia do Setor Público do Departamento de Economia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em economia.

Aprovada em 07/06/2024:

BANCA EXAMINADORA

Prof. Dr. Roberto Ferreira Tatiwa (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Leandro de Almeida Rocco
Universidade Federal do Ceará (UFC)

Prof. Dr. Jorge Eduardo Macedo Simões
Universidade Federal do Pará (UFPA)

RESUMO

A maior parte dos recursos do orçamento do estado do Ceará têm origem tributária. A exação mais importante para o erário estadual é o Imposto sobre Circulação de Mercadorias e Serviços de Transporte e de Comunicação (ICMS). A previsão dessa receita assume vital importância para execução de um orçamento público equilibrado. Portanto, visando o desenvolvimento de metodologia para aperfeiçoar a predição da receita estadual, esse estudo verifica se modelos de Vetores Autoregressivos com Fatores Dinâmicos, que incluem diversas variáveis macroeconômicas do Brasil e do Ceará, juntamente com dados da arrecadação de ICMS por CNAE, melhoram as previsões da arrecadação total do ICMS do Estado do Ceará em comparação com modelos de séries temporais do tipo ARMA. Os resultados mostram que os modelos FAVARs com o uso da CNAES melhoram as previsões da arrecadação do ICMS do Ceará quando comparados aos modelos univariados ARMA. O estudo conclui que o modelo FAVAR, usando dados da arrecadação por CNAE, apresenta melhores previsões em termos de erros quadráticos médio para a previsão da arrecadação do ICMS do estado do Ceará.

Palavras-chaves: FAVAR; ICMS; Previsão da arrecadação; CNAE;

ABSTRACT

Most of the resources in the state budget of Ceará have tax origins. The most important tax for the state treasury is the Tax on the Circulation of Goods and Transport and Communication Services (ICMS). Forecasting this revenue is vitally important for executing a balanced public budget. Therefore, aiming to develop a methodology to improve the prediction of state revenue, this study verifies whether models of Autoregressive Vectors with Dynamic Factors, which include several macroeconomic variables from Brazil and Ceará, together with data from ICMS collection by CNAE, improve the forecasts of total ICMS collection in the State of Ceará compared to ARMA-type time series models. The results show that the FAVARs models using CNAES improve forecasts of ICMS collection in Ceará when compared to the ARMA univariate models. The study concludes that the FAVAR model, using CNAE tax collection data, presents solutions in term of mean squared errors for predicting the tax collection of the best ICMS in the state of Ceará.

Keywords: FAVAR; ICMS; Revenue forecast; CNAE;

LISTA DE GRÁFICOS

Gráfico 1 - Arrecadação ICMS - Ceará.....	21
Gráfico 2 - Os quatro principais fatores estimados	24
Gráfico 3 - Previsões FAVAR (4) fora da amostra.	25

LISTA DE TABELAS

Tabela 1 - Código CNAE e descrição da atividade econômica.....	19
Tabela 2 - Variáveis utilizadas.	20
Tabela 3 - Desvio padrão, proporção da variância e proporção acumulada dos fatores.	22
Tabela 4 - MSE, MAE e MAPE dentro e fora da amostra dos modelos de previsão.....	25
Tabela 5 - Ranking de modelos.	26

SUMÁRIO

1. INTRODUÇÃO.....	8
2. REFERENCIAL TEÓRICO.....	11
2.1 Estudos correlatos de previsão de séries temporais de ICMS	11
2.2 Aplicações do Modelo FAVAR na economia	13
2.3 Modelo de Índice de Difusão (ID).....	15
3. METODOLOGIA.....	17
3.1 Modelo FAVAR	17
3.2 Base de Dados	19
4. RESULTADOS E DISCUSSÃO	22
5. CONSIDERAÇÕES FINAIS	27
REFERÊNCIAS	29

1. INTRODUÇÃO

No sistema federativo brasileiro o ente com as maiores atribuições para com a sociedade são os estados. A eles cabem a regulação e o fornecimento de um amplo leque de serviços públicos. Porém, suas fontes de receitas são restritas frente às suas obrigações¹. A maior parte de seus recursos tem origem no Imposto sobre Circulação de Mercadorias e Serviços de Transporte e de Comunicação (ICMS). O imposto representou mais de 80% das receitas tributárias estaduais arrecadas em 2021 e 2022². No estado do Ceará o imposto chegou a 40% de todas as receitas orçamentárias arrecadas no ano de 2022³.

Atualmente, os estados enfrentam um cenário de expansão em suas obrigações na prestação de serviços públicos e de restrição nas fontes de receitas, fator que pressiona, cada vez mais, o equilíbrio do orçamento público, o balanceamento entre receitas e despesas. Essa constatação demonstra que, para que possam fazer frente às suas competências e atribuições, é necessária uma acurada e precisa avaliação do seu potencial arrecadatório. Assim, é de fundamental importância a adoção de técnicas de avaliação atual e projeção futura das receitas orçamentárias, reduzindo as possibilidades de ocorrência de superestimativas e de subestimativas que irão impactar na execução orçamentária e, por conseguinte, afetar a o endividamento público e a prestação dos serviços públicos à sociedade.

Mendonça, Nascimento e Medrano (2022) ressaltam que uma acurada e precisa previsão da receita orçamentária é fundamental para fixação da despesa orçamentária e para correta execução do orçamento público, para construção de uma efetiva programação financeira e de um cronograma mensal de desembolso, e para evitar contingenciamento dos créditos e, por consequência, da realização dos gastos públicos.

Desse modo, pesquisas para o desenvolvimento de novos métodos e novas tecnologias e a procura por novos dados e informações, capazes de aumentar o poder preditivo, devem ser utilizados pelos elaboradores de políticas sempre que demonstrem que o seu uso possa reduzir o erro das estimativas entre os valores previstos e os efetivamente arrecadados. Isto posto, a presente pesquisa tenta contribuir com essa necessidade, avaliando variáveis

¹ As competências e atribuições dos estados estão previstas nos arts. 23 a 25 da Constituição Federal de 1988. Dentre essas, está competência residual, que reserva aos estados as atribuições que não lhes são vedadas pela Constituição. Dessa forma, além de outras, o que não for competência dos municípios e da União, cabem aos estados.

² Dados consultados no boletim de arrecadação dos tributos estaduais no portal do CONFAZ. Disponível em: <https://www.confaz.fazenda.gov.br/boletim-de-arrecadacao-dos-tributos-estaduais>.

³ Informe IPECE – Situação fiscal do estado do Ceará no ano de 2022.

econômicas e procurando uma modelagem econométrica que melhore e aproxime as estimativas da receita públicas.

Assim, dado seu peso na participação da receita orçamentária, a pesquisa centrará esforços para estudar um modelo de estimação da arrecadação do ICMS que um bom resultado nos indicadores de desempenho (MEA, MSE e MAPE). O modelo econométrico será focado na previsão do ICMS do estado do Ceará e utilizará os dados da arrecadação para o período de 2011 a 2022.

O presente trabalho tem por escopo verificar se os dados do ICMS do estado do Ceará, divididos por CNAEs, melhoram as previsões da arrecadação total do ICMS do Estado quando comparados com os resultados dos modelos de séries temporais ARMA. Desse modo, se busca estimar modelos de séries temporais do tipo ARMA para gerar previsões do ICMS do Estado do Ceará e com isso comparar com previsões do modelo de vetor autorregressivo aumento por fatores (FAVAR) que usam dados das arrecadações por CNAEs.

Com isso o trabalho procura responder as seguintes hipóteses: as informações dos CNAEs melhoram a previsão da arrecadação do ICMS do Estado do Ceará ao serem utilizados em um modelo de fatores quando comparadas aos de modelos autorregressivos e de médias móveis (ARMA). A inclusão de variáveis macroeconômicas do Brasil e do Ceará nos modelos de previsão aumenta a precisão das estimativas da arrecadação de ICMS do Estado do Ceará em comparação aos modelos ARMA tradicionais. O uso de dados de arrecadação de ICMS por CNAE nos modelos de Vetores Autoregressivos com Fatores Dinâmicos (FAVAR) resulta em previsões mais precisas da arrecadação total do ICMS do Estado do Ceará do que os modelos ARMA que não consideram essas variáveis.

A contribuição dessa pesquisa para literatura é avaliar a aplicação do modelo multivariado FAVAR, com componentes principais, na previsão de arrecadação do ICMS. As pesquisas anteriores sobre a arrecadação de receita de impostos estaduais buscaram medir o desempenho de modelos ARMA tendo como *benchmark* modelos estatísticos mais simples, como decomposição clássica de séries temporais e modelos de tendência (PECEGUINI, 2001) e filtros não estocásticos como os de Holt-Winters (FERREIRA, 1996); modelos ARMA (AZEVEDO, SILVA E GATSIOS, 2017) ou modelos multivariados (MOÇO, 2017) tendo como *benchmark* os modelos de projeção oficial dos estados; e, mais recentemente, o uso de modelos multivariados na comparação com modelos ARMA (MORAES, 2022). Desse modo, a pesquisa vem enriquecer essa literatura ao avaliar aplicação de modelos multivariados aumentado por fatores na previsão de receitas tributárias estaduais.

Dentre os principais resultados, pode-se destacar que o modelo FAVAR, construído com dados da arrecadação por CNAE do ICMS do Ceará e variáveis econômicas estaduais e nacionais, apresentou melhores resultados na comparação com o modelo ARMA. Verifica-se que no período de previsão de dentro da amostra o melhor modelo FAVAR reduz em até 72% o Erro Quadrático Médio de Previsão (MSE), em relação aos melhores modelos do tipo autorregressivo e de média móvel (ARMA). No período de previsão fora da amostra esse percentual é em torno de 17%.

Os resultados se coadunam com estudos anteriores, que concluíram que os modelos multivariados melhoram o desempenho da predição do ICMS do estado do Ceará quando comparados a modelos univariados (FERREIRA, 2009; MORAES, 2022) e que a utilização dos CNAEs para predição apresenta melhores resultados frente aos modelos tradicionais (MOÇO, 2017).

O presente trabalho é composto por essa introdução, e em seguida, pela segunda seção, que são abordados o referencial teórico, em que são apresenta os estudos correlatos de previsão do ICMS, o modelo utilizado e o índice de difusão. A terceira seção é composta da metodologia, em que se descreve os procedimentos executados e os dados que serão trabalhados. Na quarta seção serão discutidos os resultados das previsões e a comparação entre os modelos. Por último, na conclusão, serão apresentados os principais achados da pesquisa.

2. REFERENCIAL TEÓRICO

O referencial teórico está dividido em: apresentação de estudos correlatos da previsão da arrecadação de receitas tributárias estaduais, apresentação das aplicações modelo de FAVAR e apresentação do modelo de análise fatorial usado nesse trabalho para previsão de séries temporais.

2.1 Estudos correlatos de previsão de séries temporais de ICMS

Mendonça, Nascimento e Medrano (2022) afirmam que os trabalhos sobre arrecadação são voltados em grande parte ao estudo do nível federal, poucas são as pesquisas sobre previsão de arrecadação dos entes subnacionais. Tais estudos, são dedicados a previsões de arrecadação agregados de ICMS. Possuem como característica, o uso da modelagem Box-Jenkins, e pouco avançaram na utilização de técnicas bayesianas. Abaixo são apresentados alguns desses estudos.

Clemente e Clemente (2011) realizou um estudo de previsão da arrecadação do ICMS do Paraná para o período de agosto de 2001 a julho de 2012. Os dados utilizados para subsidiar o estudo eram referentes a janeiro de 2000 a julho de 2011. Os dados de arrecadação mensal foram modelados pela metodologia Box-Jenkins. A série foi diferenciada para torna-se não estacionária. Foi usado o método SARIMA $(2,1,0) \times (0,1,1)_{12}$ para realização da previsão. O estudo estimou que a arrecadação do estado cresceria 13,2% entre julho de 2011 e janeiro de 2012, diminuiria 5% entre janeiro e fevereiro de 2012, e aumentaria 6,7% entre fevereiro e julho de 2012.

Para o presente trabalho, é importante registrar a pesquisa de Moço (2017), que construiu modelos de previsão de arrecadação de tributária usando por base os CNAEs. Ele procedeu a elaboração de um método de previsão de arrecadação do ICMS usando modelos econométricos univariados, multivariados e com métodos de combinação tendo por objetivo a comparação com os métodos econométricos oficiais utilizados pela Secretaria da Fazenda do Rio de Janeiro, metodologia VAR com séries em nível. Usando dados de arrecadação de janeiro de 2003 a junho de 2015, ele realiza as previsões para o segundo semestre de 2015 e todo o ano de 2016. Os modelos estimados na pesquisa apresentaram melhor acurácia na previsão do que o modelo oficial. O modelo oficial apresentou erros anuais para os anos de 2015 e 2016 de, respectivamente, 14,75% e 12,64%, enquanto o modelo multivariado VEC (ICMS, IBC-BR,

PMC-RJ), que apresentou a melhor acurácia na previsão em comparação com as melhores versões de modelos univariados e de combinação, apresentou erro anual de apenas 0,74%.

Os estados têm atingido baixa acurácia na previsão de suas receitas tributárias, dado que para determinados períodos a subestimação em relação as receitas efetivamente realizadas estiveram entre 10 % e 30%. Partindo dessa constatação, Azevedo, Silva e Gatsios (2017), realizaram o seu estudo comparando a eficácia do método ARIMA frente as metodologias oficiais usadas pelos estados brasileiros. A pesquisa selecionou como amostra as previsões realizadas pelos seis estados com maior arrecadação de ICMS (SP, MG, RJ, RS, PR e BA), juntos responderam por 69,7% da arrecadação de todos os estados brasileiros no período pesquisado. As previsões do estudo foram referentes aos anos de 2012 e 2013, com base nos dados de arrecadação realizada entre 1995 e 2011. O resultado do estudo apontou para uma melhor acurácia das previsões utilizando o método ARIMA na comparação com as previsões oficiais de todos os estados pesquisados. Outra conclusão do estudo, é a crítica a literatura sobre séries temporais de receitas públicas, que se limita a apresentar modelos ideais fixos no tempo, dado que é necessário um constante repensar na utilização dos modelos utilizados frente ao surgimento de novas informações, ou seja, um modelo de previsão que apresentou uma boa acurácia para um período futuro pode, quando do surgimento de novas informações e variáveis incluídos na série, não ser o mais eficaz para realização das previsões seguintes. Assim, eles argumentam da necessidade do treinamento especializados dos quadros técnicos para que possam manejar vários métodos e novas informações e, com isso, construir previsões mais acuradas e com maior qualidade.

Por fim, ao analisarem vários estudos sobre previsão de receitas estaduais, Mendonça, Nascimento e Medrano (2022) observam que poucos estudos desenvolveram previsões de arrecadação para o médio prazo. Não obstante, esses estudos condicionaram suas avaliações a cenários macroeconômicos de médio prazo.

Com relação avaliação da utilização de técnicas de aprendizagem estatística para previsão de variáveis é mister comentar o artigo de Barbosa, Ferreira e Da Silva (2020) que avalia se a utilização dessas técnicas melhora as previsões de importantes variáveis macroeconômicas da economia brasileira. A pesquisa observou o desempenho de modelos fatoriais combinados com uso de técnicas de aprendizagem – métodos de *shrinkage*, combinações de previsões e seleção de previsores – na previsão das variáveis macroeconômicas taxa de desemprego, índice de produção industrial, IPC e IPCA. O estudo concluiu que os modelos que combinaram técnicas de supervisão fatorial e aprendizagem estatística apresentaram melhor desempenho na previsão das variáveis macroeconômicas em comparação

a modelos sem fatores, a modelos fatoriais com o sem supervisão e a modelos de aprendizado sem fatores.

No que se refere a previsão de ICMS do estado do Ceará, Ferreira (1996) mostra que modelos ARIMA geram as previsões mais eficientes do que filtros não estocásticos como os de Holt-Winters e que a técnica de combinação de previsões desses modelos pode melhorar a previsão dessa variável.

Rodrigues de Moraes (2022) avalia o desempenho de modelos univariados e multivariados na estimação da arrecadação dessa receita. O estudo utilizou os dados da arrecadação do imposto entre os anos de 2003 e 2019 em frequência mensal. As previsões foram realizadas sobre dois horizontes temporais, a curto prazo (12 meses) e a médio prazo (48 meses). As variáveis usadas na pesquisa foram a arrecadação do ICMS, o Índice de Vendas de Varejo (IVV), a Demanda Turística (DT), o Índice de Produção Industrial (IPI) e o Índice de Atividade Econômica do Banco Central para o Estado do Ceará (IBC-CE). A métrica usada na pesquisa como indicador de qualidade da previsão foi o RMSE (Raiz Quadrada dos Erros Quadráticos Médios). O modelo univariado que apresentou melhor previsão foi o ARIMA (1,1,1) (1,1)12, que apresentou o RSME de 5,05%, no curto prazo, e de 4,95%, no médio prazo. O modelo multivariado usado na pesquisa foi o Vetor Autorregressivo (12), que apresentou o melhor desempenho quando usou como proxy a variável IBC-CE, alcançando o RSME de 3,99%, no curto prazo, e de 3,60%, no médio prazo. Assim, o estudo concluiu que o modelo multivariado apresentou melhor precisão na sua previsão da arrecadação do ICMS em comparação as previsões dos modelos univariados.

Buscando investigar a existência e as características de ciclos econômicos na taxa de crescimento do ICMS do estado do Ceará, Ferreira (2009) analisa a eficácia da previsão de modelos de regressão linear e de modelos de efeito *threshold* e com mudança de regimes Markovianas, tendo como benchmark os modelos ARMA. Usando dados da arrecadação de junho/2005 a maio/2006 para predição de junho/2006, o autor reporta que o modelo multivariado de efeito *threshold* apresentaram os menores de Erros Quadráticos Médios (EQM), melhorando a eficiência preditiva dos modelos ARMA lineares.

2.2 Aplicações do Modelo FAVAR na economia

O modelo FAVAR é aplicado em várias áreas da ciência econômica como um preditor que oferece bons resultados em comparação a outros modelos tradicionais. Em pesquisa sobre estrutura temporal da curva de juros no Brasil, Vieira, Chague e Fernandes

(2017), utilizam uma metodologia VAR aumentada por fatores Nelson-Siegel para análise e previsão do comportamento dessa variável. Os modelos disponíveis, normalmente, em comparação ao passeio aleatório de referência, só apresentam resultados superiores em horizontes temporais acima de seis meses. A pesquisa demonstrou, que o modelo FAVAR no horizonte de um mês, teve, em média, erros de previsão de 5% a 10% inferiores ao modelo de passeio aleatório de referência. Para o horizonte de dois meses, o modelo gera previsões com o erro absoluto médio de 30 % a 40% inferiores ao produzido pelo passeio aleatório. Os resultados encontrados demonstram que o uso dessa metodologia melhora os resultados preditivos no curto prazo, superando o benchmark, e tem bons resultados para horizontes maiores.

O modelo FAVAR tem como uma de suas vantagens a possibilidade de se trabalhar com grandes volumes de dados. Madhou, Sewak, Mossa e Ramiah (2019), afirmam que os modelos econométricos tradicionais não são projetados para se trabalhar com um grande número de variáveis, o que impõe ao modelador a tarefa de ter que escolher entre conjuntos de variáveis complexos. Assim, Madhou, Sewak, Mossa e Ramiah (2019), elaboraram um modelo FAVAR, usando diversas séries de variáveis (índices de preços, taxas de câmbio, índices de atividade e outros) com o objetivo de estimar o crescimento do PIB real de uma pequena economia aberta. O estudo concluiu que o modelo FAVAR pode melhorar significativamente a previsão quando a pesquisa envolver um grande conjunto de dados. O modelo FAVAR gerou melhores resultados na comparação ao modelo BVAR. Estudo elaborado por Paccaginini (2017), vai na mesma linha, ao demonstrar que o modelo FAVAR apresentou melhor desempenho para previsão das principais variáveis da economia americana na comparação com modelos clássicos e com o modelo VAR bayesiano. Ademais, o estudo demonstrou que a inclusão de fatores em um modelo VAR aumenta o poder preditivo sobre as principais variáveis da economia americana.

Na comparação entre os modelos FAVAR e os modelos tradicionais AR, Figueredo e Guillén (2013), avaliam o desempenho da acuraria preditiva FAVAR para prever a inflação de preços ao consumidor no Brasil. Os autores concluíram que em um horizonte de previsão de longo prazo da inflação os modelos FAVAR apresentam melhor desempenho em comparação aos modelos AR. Figueredo e Guillén (2013) afirmam que ao se trabalhar com previsões é importante o foco na escolha das variáveis alvo, que tem um impacto sobre o comportamento futuro da série, para gerar boas previsões medianas.

2.3 Modelo de Índice de Difusão (ID)

Assim, o presente trabalho utiliza o modelo FAVAR de índice de difusão (ID), técnica que se aproxima de um modelo de análise fatorial dinâmico. Esse é um modelo parcimonioso (FERREIRA, 2005), princípio que afirma que o melhor modelo deve usar a menor quantidade de variáveis possíveis sem prejudicar seu desempenho, ou seja, variáveis não devem ser incluídas sem necessidade.

Como um modelo de análise fatorial, o ID é usado para pesquisas nas quais existam uma grande quantidade de variáveis na qual dificulte sua manipulação e análise. Segundo Matos e Rodrigues (2019) os modelos de análise fatorial têm como principais usos o resumo e a redução dos dados, sendo úteis quando a pesquisa pretende recorrer a técnicas multivariadas com grande número de variáveis.

O DI por ser um modelo fatorial é um modelo de variável latente. Bartholomew e Knott (1999) afirmam que entre as variações de um fenômeno estudado existem variáveis observáveis e não observáveis. As não observáveis, que são percebidas indiretamente por movimento comuns entre outras variáveis que promovem variação no fenômeno são denominadas de variável latente. Essas não podem ser tratadas diretamente, só sendo possível analisá-las por meio dos movimentos comuns entre outras variáveis. Assim, sendo f a quantidade de variáveis não observáveis representado por r e sendo x a quantidade de variáveis observáveis representadas por k , temos que $x < k$.

Diferente do modelo de análise fatorial dinâmico tradicional que utiliza a estimativa de máxima verossimilhança (MLE) para reduzir a dimensionalidade dos dados, os índices de difusão usam a estimativa de análise de componentes principais (FERREIRA; BIERENS; CASTELAR, 2005). Segundo Stock e Watson (1998), a análise de componentes principais quando usado para estimar fatores tem como vantagens a capacidade de se trabalhar com um grande número de preditores e a facilidade de serem generalizados quando da irregularidade dos dados. Ferreira (2005) afirma que a estimativa por componentes principais gera fatores que mantêm a sua consistência mesmo quando o número de variáveis tende ao infinito, sendo uma característica importante para um trabalho empírico que conte com um número considerável de variáveis as quais tenham apenas algumas observações sobre elas.

Os índices de difusão, segundo Stock e Watson (2002), medem um movimento comum em um conjunto de variáveis macroeconômicas. O estudo realizado pelos autores teve por objetivo a construção de séries temporais macroeconômicas usando 215 variáveis que refletissem a economia americana. Usando um modelo de componente principal foram

estimados fatores, que reduziram os diversos preditores em poucos fatores, construindo os índices de difusão e, em seguida, promoveu as previsões. O estudo usou como benchmark os modelos autorregressivos (AR), indicadores antecedentes multivariados (LI) e vetor autorregressivo (VAR). A pesquisa encontrou apenas seis fatores que explicavam a maior parte das variâncias das 215 séries temporais usadas e que alguns fatores, no caso americano, foram necessários para previsão da atividade real da economia. Assim, o estudo sugeriu que um vetor de estado pequeno deve ser necessário para previsão de séries macroeconômicas. Stock e Watson (2002) concluíram que os modelos de índices de difusão apresentam resultados encorajadores.

3. METODOLOGIA

3.1 Modelo FAVAR

Seguimos o procedimento metodológico proposto por Bernanke, Boivin e Elias (2005). Seja Y_t , um vetor $M \times 1$ de variáveis macroeconômicas observáveis, e F_t , um vetor $K \times 1$ de fatores não observados que contém informações adicionais. A dinâmica conjunta (F_t', Y_t') é dada por:

$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \Phi(L) \begin{bmatrix} F_{t-1} \\ Y_{t-1} \end{bmatrix} + v_t \quad (1)$$

onde $\Phi(L)$ é o polinômio de defasagens de ordem finita d . E $v_t(0, Q)$. O modelo de vetor autorregressivo aumentado por fatores (FAVAR) aninha análises VAR padrão, então a estimativa da equação (1) permite comparações com os resultados VAR existentes e permite avaliar a contribuição marginal da informação adicional contida em F_t .

Seja X_t um vetor $N \times 1$ de variáveis informacionais relacionadas aos fatores não observáveis F_t e às variáveis observadas Y_t por uma equação da forma:

$$X_t = \Lambda^f F_t + \Lambda^y Y_t + e_t \quad (2)$$

onde Λ^f é uma matriz $N \times K$ de cargas fatoriais, Λ^y é uma matriz $N \times M$ e $e_t N(0, W)$. Os termos de erro podem ser não correlacionados ou fracamente correlacionados, dependendo se a estimativa é feita por métodos de verossimilhança ou componentes principais. v_t e e_t são independentes.

Os autores consideram duas abordagens para estimar (1) e (2). A primeira, utiliza componentes principais em duas etapas e a segunda é uma abordagem de probabilidade bayesiana de etapa única. Estas abordagens diferem em várias dimensões e não é claro a priori que uma deva ser favorecida em detrimento da outra. A primeira abordagem é utilizada neste trabalho. Primeiro, o espaço abrangido pelos fatores é estimado usando os primeiros componentes principais $K + M$ de X_t ; em seguida, o FAVAR, equação (1), é estimado por métodos padrão, com F_t substituído por \hat{F}_t . A abordagem semiparamétrica é computacionalmente mais simples, embora não explore a estrutura da equação de transição na estimação dos fatores.

Como benchmark, utilizam-se modelos univariados autorregressivos de média móvel (ARMA) para realizar previsões dentro e fora da amostra. Seja y_t uma série temporal que segue um processo ARMA(p, q) como:

$$y_t = \sum_{i=1}^p \phi_i x_{t-i} + \sum_{j=0}^q \theta_j \epsilon_{t-j} \quad (3)$$

onde p e q representam, respectivamente, o número de defasagens do modelo AR e MA. Para verificar a hipótese de estacionariedade da série x_t são conduzidos testes de estacionariedade, usando o teste de Dickey-Fuller Aumentado (ADF), proposto por Dickey e Fuller (1979) em que a hipótese nula é a presença de uma raiz unitária.

Por fim, realizam-se as previsões dentro (*in-sample*) e fora da amostra (*out-of-sample*). As previsões *in-sample*, são os *fitted values* do modelo estimado. Para calcular as previsões *out-of-sample*, deve-se dividir o conjunto de dados em amostras de treino (T) e teste (H). Na primeira, estima-se o modelo e na segunda o realizam-se as previsões. A acurácia dessas previsões é avaliada pelas seguintes estatísticas:

$$MSE_{t,H} = \frac{\sum_{h=1}^H e_t(h)^2}{H}$$

$$MAE_{t,H} = \frac{\sum_{h=1}^H |e_t(h)|}{H} \quad (4)$$

$$MAPE_{t,H} = \sum_{h=1}^H \left| \frac{e_t(h)}{y_{t+h}} \right|$$

onde as siglas MSE, MAE e MAPE significam, respectivamente, o erro quadrático médio, o erro absoluto médio e o erro percentual absoluto médio. O MSE penaliza mais severamente grandes erros devido à elevação ao quadrado das diferenças. É mais usada em contextos quando se deseja minimizar grandes desvios. O MAE e o MAPE usam valores absolutos do erro de previsão, por isso não penalizam grandes erros como o MSE.

3.2 Base de Dados

O objetivo desse trabalho é realizar previsões para a taxa de crescimento da arrecadação do Imposto sobre Circulação de Mercadorias e Serviços (ICMS) do Ceará por meio de vetores autorregressivos aumentados por fatores (FAVAR). Para isso, utiliza-se um conjunto de séries de arrecadação por categoria de atividade e séries econômicas.

As séries de arrecadação do ICMS compreendem os 20 tipos de atividades econômicas, de acordo com a Classificação Nacional de Atividades Econômicas (CNAE), de maior participação na arrecadação entre 2011.1 e 2022.12. A opção pela quantidade de vinte CNAEs ocorreu em virtude de que individualmente apresentaram arrecadação média superior a 1% e de que combinados representaram mais de 61% da arrecadação do ICMS do período. Os setores de maior peso estão relacionados à fabricação e/ou comércio de energia, bebidas, telefonia, alimentos, medicamentos e automóveis.

Tabela 1 - Código CNAE e descrição da atividade econômica

<i>Código</i>	<i>Descrição</i>
1921700	Fabricação de produtos do refino de petróleo.
3514000	Distribuição de energia elétrica.
1113502	Fabricação de cervejas e chopes.
4681801	Comércio atacadista de álcool carburante, biodiesel, gasolina e demais derivados de petróleo, exceto lubrificantes.
6120501	Telefonia móvel celular.
4639701	Comércio atacadista de produtos alimentícios em geral.
4644301	Comércio atacadista de medicamentos e drogas de uso humano.
3511501	Geração de energia elétrica.
6110801	Serviços de telefonia fixa comutada.
2910701	Fabricação de automóveis, camionetas e utilitários.
4781400	Comércio varejista de artigos do vestuário e acessórios.
9999990	Receita pessoa jurídica sem CGF
1122401	Fabricação de refrigerantes.
3520401	Produção de gás; processamento de gás natural.
1122401	Produção de preparações em pó ou em xarope para fabricação de refrigerantes, para fins industriais.
4646001	Comércio atacadista de cosméticos e produtos de perfumaria.
4635402	Comércio atacadista de cerveja, chope e refrigerante.
4753900	Comércio varejista especializado de eletrodomésticos e equipamentos de áudio e vídeo.
4711302	Comércio varejista de mercadorias em geral, com predominância de produtos alimentícios - supermercados.
4636202	Comércio atacadista de cigarros, cigarrilhas e charutos.

Fonte: Sefaz/CE.

Nota: atividade econômicas organizadas em ordem decrescente de participação na arrecadação de ICMS.

A tabela 2 abaixo apresenta o conjunto de séries econômicas utilizadas. Os dados possuem periodicidade mensal e compreendem o período de 2011.1 a 2022.12. Todas as séries correspondem ao estado do Ceará, exceto a taxa de inflação para a Região Metropolitana de Fortaleza (RMF), a taxa de juros Selic e a taxa de câmbio.

As séries de arrecadação e econômicas totalizam 33 variáveis. Aquelas mensuradas em reais (R\$) foram deflacionadas pelo IPCA a preços de dezembro de 2022. Os testes de estacionariedade foram conduzidos e aquelas que possuem uma raiz unitária tomou-se suas primeiras diferenças.

O gráfico 1, mostra a série real de arrecadação do ICMS do Ceará. A série apresenta um forte comportamento sazonal, com aumento da arrecadação entre os meses finais e iniciais do ano, decorrente da maior movimentação econômica nesses meses. A arrecadação tem um comportamento pró-cíclico com a atividade econômica, cresce de 2011 a 2013 e de 2017 a 2022 e decresce de 2014 a 2016.

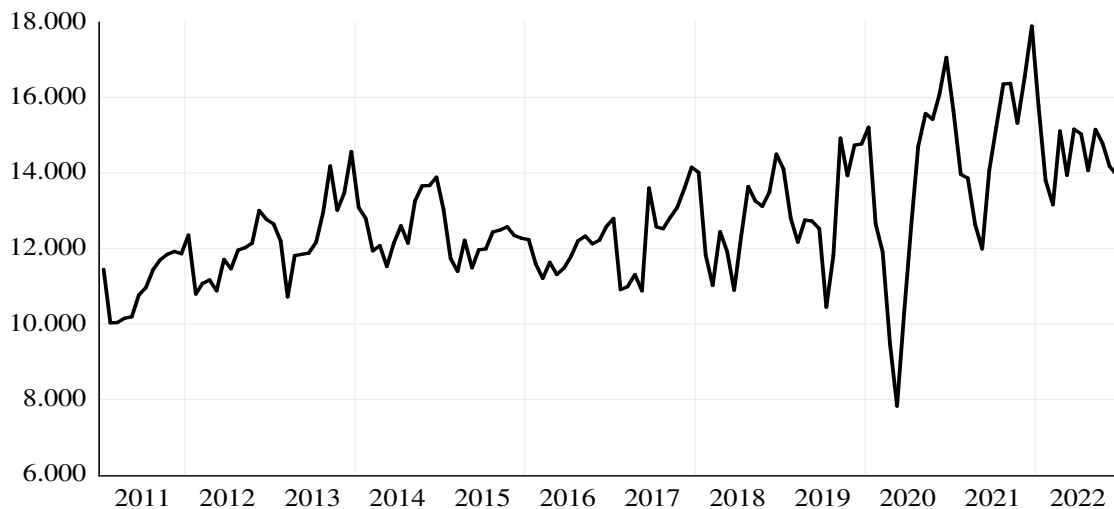
Tabela 2 - Variáveis utilizadas.

<i>Variável</i>	<i>Descrição</i>	<i>Medida</i>	<i>Fonte</i>
ICMS	Arrecadação do ICMS	R\$	Ipea
Inflação	Taxa de Inflação IPCA - RMF	%	IBGE
PIB	Produto Interno Bruto	R\$	IPECE
Juros	Taxa de juros Selic	%	BACEN
Câmbio	Taxa de câmbio R\$/US\$	%	BACEN
Energia	Consumo de eletricidade	MWh	Aneel
Demissões	Número de demissões	Unidade	RAIS/CAGED
Admissões	Número de admissões	Unidade	RAIS/CAGED
Trabalho	Saldo do trabalho	Unidade	RAIS/CAGED
Etanol	Preço médio do etanol	R\$	ANP
Gasolina	Preço médio da gasolina	R\$	ANP
Diesel	Preço médio do diesel	R\$	ANP
Combustíveis	Preço médio dos combustíveis	R\$	ANP

Fonte: elaboração própria.

Nota: A Região Metropolitana de Fortaleza (RMF) compreende dezenove municípios, incluindo a capital do estado.

Gráfico 1 – Arrecadação ICMS - Ceará.



Fonte: elaboração própria

Nota: valores reais a preços de dezembro/2022.

Entre os meses de abril e maio de 2020 ocorre a maior queda na arrecadação, acumulada em 37,64%, devido à baixa atividade econômica proveniente das medidas de isolamento durante a pandemia de Covid-19.

Após o impacto inicial, com o relaxamento das medidas de restrição social, arrecadação do imposto teve uma recuperação. O ano de 2021, segue dinâmica semelhante ao ano anterior, queda nos primeiros meses e recuperação nos seguintes. Desse modo, pode-se inferir que, nos dois anos iniciais da pandemia, houve quedas na arrecadação no início do exercício que foi compensado pelo excepcional crescimento no segundo semestre, momento em que as medidas de isolamento social foram relaxadas e que houve um aumento da arrecadação em virtude das políticas de transferência de renda do Governo Federal (IPECE, 2022; MORAES et al., 2023).

O ano 2022, a série da arrecadação apresenta alteração no comportamento, há redução da volatilidade. O comportamento da arrecadação no exercício, apresenta comportamento diverso do demonstrado em anos anteriores, 2017; 2018 e 2019, a eclosão da pandemia de Covid-19, constatação que pode evidenciar a alteração da dinâmica de arrecadação do ICMS no Ceará (IPECE, 2022).

4. RESULTADOS E DISCUSSÃO

Como apresentado anteriormente, os dados possuem periodicidade mensal e compreendem o período de 2011.1 a 2022.12. Inicialmente, testa-se a hipótese nula de raiz unitária com o teste ADF. As séries com raiz unitária foram diferenciadas até se obter variáveis estacionárias. Nos modelos fatoriais, vale lembrar que a matriz de dados dos previsores é padronizada, para se utilizar a técnica dos modelos de fatores em comum. Para todos os modelos, período de treino, previsão dentro da amostra (in sample), compreende os meses entre 2011.1 a 2020.06. O período de teste, previsão fora da amostra (out-of-sample) usa os meses entre 2020.07 a 2022.12 (30 meses).

O modelo de análise fatorial baseado em índices de difusão foi empregado e gerou um máximo de dez fatores que foram estimados por análise de componentes principais. A análise fatorial foi escolhida para reduzir a dimensionalidade dos dados e identificar os fatores subjacentes que explicam a variabilidade na arrecadação do ICMS. Este método é particularmente útil quando se trabalha com um grande número de variáveis correlacionadas, como no caso dessa dissertação, onde há 33 variáveis. Dessa forma utilizou-se a Análise de Componentes Principais (PCA) para extrair os fatores. Foram selecionados 10 componentes principais, que explicam a quase totalidade da variação dos dados.

Assim, visando reduzir as 33 variáveis originais, foram gerados os componentes presentes na Tabela 3, que explicam a quase totalidade variação dos dados. Vale frisar, que os componentes não são correlacionados entre si.

Tabela 3 - Desvio padrão, proporção da variância e proporção acumulada dos fatores.

	<i>Desvio Padrão</i>	<i>Proporção da Variância</i>	<i>Proporção Acumulada</i>
PC1	40950000	53,50%	53,50%
PC2	27960000	24,94%	78,44%
PC3	11660000	4,34%	82,77%
PC4	11410000	4,15%	86,93%
PC5	9333000	2,78%	89,71%
PC6	8034000	2,06%	91,76%
PC7	7701000	1,89%	93,66%
PC8	6243000	1,24%	94,90%
PC9	5061000	0,82%	95,72%
PC10	4804000	0,74%	96,45%

Fonte: elaboração própria

Nota: máximo de 10 componentes principais empregados.

Os resultados mostram que o primeiro componente principal (PC1) explica 53,50% da variância total dos dados. Isso indica que mais da metade da informação contida nos dados originais pode ser resumida neste único componente. O segundo componente principal (PC2) explica 24,94% da variabilidade total, e, juntamente com o PC1, os dois primeiros componentes explicam 78,44% da variabilidade dos dados, o que representa uma grande parte da informação contida nos dados originais.

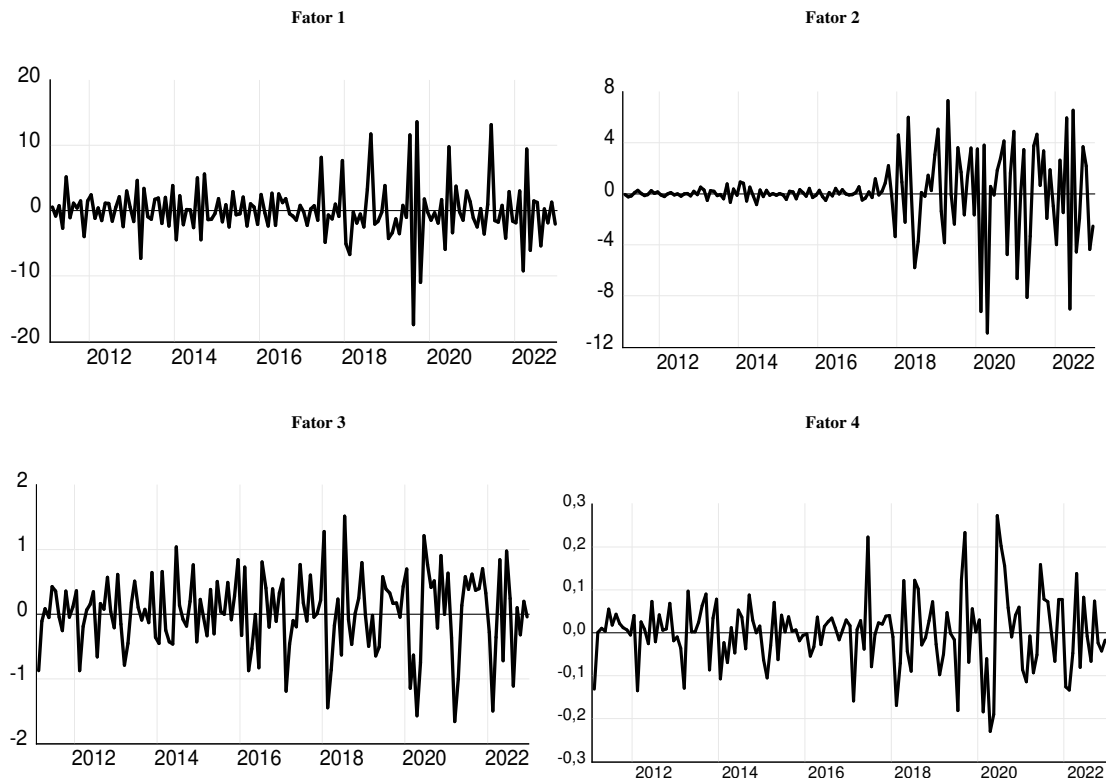
Os componentes principais de PC3 a PC10 explicam porções progressivamente menores da variabilidade total, com o PC10 explicando apenas 0,74% da variância. No entanto, a proporção acumulada de variância explicada pelos dez primeiros componentes principais é de 96,45%, indicando que a maior parte da variabilidade pode ser capturada por esses dez componentes.

O uso dos componentes principais (PCA), em vez das variáveis originais, revela-se uma alternativa para a redução da dimensionalidade dos dados. Desta forma, essa metodologia pode ajudar a gerar modelos que atendem o “princípio da parcimônia” (FERREIRA, 2005) e, portanto, capazes de gerar previsões mais eficientes.

Os quatro primeiros componentes principais (PC1 a PC4) juntos explicam 86,93% da variabilidade total dos dados. Isso significa que quase 87% da informação contida nos dados originais pode ser capturada utilizando apenas esses quatro componentes. Este nível de explicação é consideravelmente alto, indicando que os componentes selecionados retêm a maior parte da variância original dos dados.

Utilizar apenas os quatro primeiros componentes principais permite uma redução significativa na dimensionalidade dos dados, passando de 33 variáveis originais para apenas 4 componentes. Isso simplifica a modelagem e a análise subsequente, tornando o processo mais manejável e eficiente, sem perder uma quantidade substancial de informação. Os componentes principais a partir do quinto (PC5 e seguintes) explicam proporções muito menores da variância total (menos de 3% cada). Ao excluir esses componentes menos relevantes, reduzimos o risco de incluir ruído e variabilidade insignificante na modelagem, o que pode melhorar a robustez e a precisão das previsões.

Gráfico 2 - Os quatro principais fatores estimados



Fonte: elaboração própria

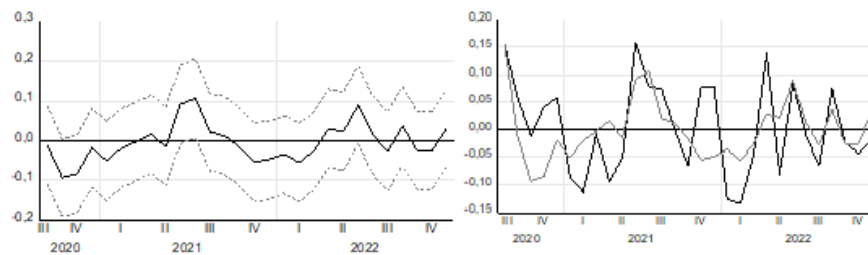
Nota: fatores estimadores por PCA.

Ao utilizar os quatro primeiros componentes principais na modelagem do FAVAR, garante-se que a maioria da variabilidade dos dados é capturada de forma eficiente e interpretável. Isso não só melhora a precisão das previsões, mas também simplifica a análise, tornando-a mais acessível e acionável para tomadores de decisão e gestores de políticas públicas. A alta proporção da variância explicada por esses componentes justifica plenamente sua utilização exclusiva na análise. A dinâmica dos quatro principais fatores estimados se encontram no gráfico 2.

A tabela 6 mostra o MSE, o MAE e o MAPE para a análise de previsão no período de treino (in sample) e no período de teste ou previsão fora da amostra. No primeiro período, o modelo FAVAR com 12 defasagens apresenta os menores MSE e MAE. O modelo AR (1) obteve o menor valor de MAPE. Como explicado na seção sobre metodologia, o MSE penaliza grandes erros de previsão ao elevá-los ao quadrado, enquanto MAE e MAPE não o fazem ao usarem os valores absolutos dos mesmos. Desta forma, opta-se por concentrar a análise no MSE, uma vez que grandes erros de previsão de receita podem ter efeitos mais negativos para a administração pública. O MSE do FAVAR(12) é 72% menor do que o do melhor modelo ARMA para esse período de previsão, que foi o AR(1).

Ao se analisar o período de previsão de fora da amostra, os modelos FAVAR(2) e FAVAR(4) apresentam os menores MSE. Nesse caso, pode-se optar pelo FAVAR(2) em virtude do princípio da parcimônia, ou pelo FAVAR(4) que apresenta mesmo MSE e menores MAE e MAPE. O gráfico 4 apresenta os valores reais e as previsões para esse período do FAVAR(4).

Gráfico 3 - Previsões FAVAR (4) fora da amostra.



Fonte: elaboração própria

Nota: a linha preta representa as previsões e a linha cinza os erros padrão.

Nesse período de previsão o modelo ARMA com menor MSE foi o MA(1). Vale ressaltar que o modelo AR(1) gera previsões muito similares a do MA(1) em termos de MSE e MAE, mas com um MAPE menor. De acordo com a comparação entre os MSEs da previsão fora da amostra, o FAVAR(4) possui MSE cerca de 17% menor do que os modelos AR(1) e MA(1).

Tabela 4 - MSE, MAE e MAPE dentro e fora da amostra dos modelos de previsão.

	<i>in sample</i>			<i>out of sample</i>		
	MSE	MAE	MAPE	MSE	MAE	MAPE
AR(1)	0,006453	0,055597	111,6004*	0,006219	0,067531	116,9714*
AR(2)	0,006348	0,056247	157,7949	0,006079	0,068928	140,4695
AR(3)	0,006394	0,056847	152,5564	0,006211	0,070228	169,9270
AR(4)	0,006152	0,058035	210,9262	0,006476	0,070779	166,8423
MA(1)	0,006544	0,056397	119,3179	0,006198	0,067297	121,3555
MA(2)	0,005702	0,056796	395,8840	0,009604	0,081329	515,2749
MA(3)	0,005460	0,054802	284,2504	0,008218	0,071819	498,7366
MA(4)	0,005027	0,052854	257,8945	0,006466	0,064659	306,8170
ARMA(1,1)	0,005499	0,052627	221,8455	0,006470	0,066894	365,9883
ARMA(1,2)	0,005161	0,053016	250,2884	0,006534	0,064170	358,6785
ARMA(1,3)	0,005160	0,050140	248,0610	0,006556	0,064341	361,9752
ARMA(1,4)	0,004916	0,052148	259,1379	0,006461	0,064636	306,1806
ARMA(2,1)	0,005156	0,053242	234,9918	0,006952	0,065981	391,5465
ARMA(2,2)	0,005106	0,053014	263,1600	0,006234	0,064405	314,9330
ARMA(2,3)	0,005067	0,053190	261,4915	0,006338	0,064325	326,4311

ARMA(2,4)	0,004957	0,052570	250,9888	0,006413	0,064752	313,5530
ARMA(3,1)	0,005160	0,053142	220,0376	0,007287	0,067646	403,9043
ARMA(3,2)	0,071105	0,053449	263,3646	0,006612	0,064572	343,7017
ARMA(3,3)	0,004787	0,051910	245,5341	0,008141	0,067814	373,7215
ARMA(3,4)	0,004326	0,049442	263,2791	0,008556	0,070363	239,0937
ARMA(4,1)	0,004829	0,051419	252,8464	0,008343	0,066461	264,4686
ARMA(4,2)	0,004835	0,051831	243,8188	0,007370	0,065330	308,5269
ARMA(4,3)	0,004674	0,051045	247,5937	0,008682	0,068111	312,4320
ARMA(4,4)	0,004029	0,047435	232,8911	0,008996	0,074210	280,5164
FAVAR(1)	0,006137	0,053622	161,7057	0,006139	0,063424	159,1604
FAVAR(2)	0,005899	0,053282	161,7513	0,005154*	0,060779	167,4137
FAVAR(3)	0,005916	0,053639	162,6139	0,005402	0,063851	237,5775
FAVAR(4)	0,004853	0,052639	215,9683	0,005154*	0,059911*	160,6118
FAVAR(5)	0,004749	0,052007	221,6157	0,005639	0,061444	207,9310
FAVAR(6)	0,004156	0,046915	238,5544	0,008256	0,070184	367,2608
FAVAR(7)	0,003941	0,049282	248,7536	0,010705	0,080173	552,8996
FAVAR(8)	0,002961	0,042439	230,5813	0,011071	0,078158	366,5669
FAVAR(9)	0,002523	0,038735	242,2820	0,016360	0,101399	456,5195
FAVAR(10)	0,002343	0,037622	274,4182	0,023852	0,119229	607,1569
FAVAR(11)	0,002262	0,036251	283,2023	0,028799	0,131891	820,4419
FAVAR(12)	0,001792*	0,033482*	313,2689	0,037462	0,156318	1233,8900

Fonte: elaboração própria.

Nota: * é o menor valor do MSE ou MAE ou MAPE.

A tabela 5, mostra o ordenamento dos 3 melhores modelos com menores MSE, MAE e MAPE. Verifica-se que os modelos FAVAR apresentam menores MSE e MAE, tanto no período de previsão de dentro da amostra quanto no de fora da amostra.

Tabela 5 - Ranking de modelos.

	<i>in sampe</i>			<i>out of sample</i>		
	MSE	MAE	MAPE	MSE	MAE	MAPE
1°	FAVAR(12)	FAVAR(12)	AR(1)	FAVAR(2)	FAVAR(4)	AR(1)
2°	FAVAR(11)	FAVAR(11)	MA (1)	FAVAR(4)	FAVAR(2)	MA (1)
3°	FAVAR(10)	FAVAR(10)	AR(3)	FAVAR(3)	FAVAR(5)	AR(2)

Fonte: elaboração própria.

5. CONSIDERAÇÕES FINAIS

O presente trabalho teve como principal objetivo verificar se modelos vetoriais com fatores em comum (FAVAR) podem melhorar as previsões de modelos autorregressivos e de média móvel para a taxa de crescimento da arrecadação do ICMS no estado do Ceará.

Os resultados demonstram que o modelo FAVAR se mostrou uma ferramenta robusta e eficaz para prever a arrecadação do ICMS no Ceará. Ao incorporar fatores latentes que capturam relações complexas entre variáveis econômicas, o FAVAR permitiu uma compreensão mais profunda dos determinantes da receita fiscal estadual.

Verifica-se que no período de previsão de dentro da amostra o melhor modelo FAVAR reduz em até 72% o Erro Quadrático Médio de Previsão (MSE), em relação aos melhores modelos do tipo autorregressivo e de média móvel (ARMA). No período de previsão fora da amostra esse percentual é em torno de 17%.

Ao comparar os resultados encontrados com estudos anteriores na literatura⁴, observamos consistência nas vantagens do modelo FAVAR em capturar relações complexas entre variáveis econômicas, proporcionando uma visão mais holística e dinâmica da arrecadação fiscal. Essa comparação valida a robustez do modelo FAVAR e destaca sua relevância como uma ferramenta poderosa para previsão econômica regional, complementando e expandindo o conhecimento existente na área.

Além dos benefícios práticos, é crucial reconhecer as limitações inerentes ao estudo. Os modelos de previsão, incluindo o FAVAR, dependem de suposições e da qualidade dos dados utilizados, o que pode impactar a validade das previsões. Discutir essas limitações não apenas contextualiza os resultados, mas também identifica áreas para futuras melhorias e pesquisas. Explorar técnicas não lineares que permitem a mudança dos parâmetros de acordo com diferentes regimes econômicos (como em Ferreira, 2009) e integrar novos indicadores econômicos podem aprimorar ainda mais a capacidade preditiva do FAVAR e expandir seu potencial de aplicação em diferentes cenários econômicos e políticos.

Dessa forma, para estudos futuros, recomenda-se explorar formas funcionais diferentes, que incluam possíveis não linearidades na dinâmica do ICMS, bem como a análise da eficiência preditiva de modelos de aprendizado de máquina (machine learning) e de aprendizado profundo (deep learning).

⁴ Pessoa, Coronel e Lima (2013); Bernadino et al. (2021)

Por fim, essa dissertação contribui significativamente para o campo da previsão econômica regional ao demonstrar o potencial do modelo FAVAR na modelagem e previsão da arrecadação do ICMS no Ceará. Os resultados apresentados não apenas corroboram a eficácia do FAVAR, mas também delineiam um caminho promissor para futuras investigações e aplicações práticas na gestão econômica estadual.

REFERÊNCIAS

- AZEVEDO, Ricardo Rocha de e SILVA, José Marcos da e GATSIOS, Rafael Confetti. Análise crítica dos modelos de previsão de série temporal com base no ICMS estadual. **Revista de Gestão, Finanças e Contabilidade**, v. 7, n. 1, p. 164-184, 2017. Disponível em: https://www.researchgate.net/publication/312542601_Analise_Critica_dos_Modelos_de_Previao_de_Serie_Temporal_com_Base_no_ICMS_Estadual. Acesso em: 30 set. 2023.
- BRASIL. Lei complementar nº 101, de 04 de maio de 2000. Estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal e dá outras providências. Lei de Responsabilidade Fiscal. – 3. ed. **Coordenação de Edições Técnicas do Senado Federal**. Brasília: Senado Federal, 2022. Disponível em: https://www2.senado.leg.br/bdsf/bitstream/handle/id/602080/Lei_responsabilidade_fiscal_3ed.pdf . Acessado em: 28 set. 2023.
- BARBOSA, Rafael B.; FERREIRA, Roberto Tatiwa; DA SILVA, Thibério Mota. Previsão de variáveis macroeconômicas brasileiras usando modelos de séries temporais de Alta Dimensão. **Revista de Estudos Econômicas** (USP). v. 50, n.1, p. 67-98, 2020. Disponível em: <https://www.scielo.br/j/ee/a/7ZSSwmy7spqjDyYWR4NPCvP/?lang=pt> . Acesso em: 05. fev. 2024.
- BARTHOLOMEW, D. J.; KNOTT, M. **Latent Variable models and factor analysis**. New York: Oxford University Press Inc, 1999.
- BERNARDINO, Wilton et al. PREVISÃO DA ARRECADAÇÃO DE ICMS PARA O ESTADO DE SERGIPE: MODELAGEM A PARTIR DA METODOLOGIA DE BOX E JENKINS. **Latin American Journal of Business Management**, v. 12, n. 1, 2021
- BERNANKE, Ben S.; BOIVIN, Jean; ELIASZ, Piotr. Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. **The Quarterly journal of economics**, v. 120, n. 1, p. 387-422, 2005.
- BUENO, Rodrigo de Losso da Silveira. **Econometria de Séries Temporais**. 2 ed. São Paulo: Cengage, 2020.
- CEARÁ (Estado). Instituto de Pesquisa e Estratégia Econômica do Ceará – IPECE. **Informe Nº 228 – Situação Fiscal do Estado do Ceará no ano de 2022**. Disponível em: https://www.ipece.ce.gov.br/wp-content/uploads/sites/45/2023/05/ipece_informe_228_12Mai_2023.pdf . Acesso em: 18. set. 2023.
- CEARÁ (Estado). Instituto de Pesquisa e Estratégia Econômica do Ceará – IPECE. **Informe Nº 239 – Comércio exterior do Ceará em 2023**. Disponível em: https://www.ipece.ce.gov.br/wp-content/uploads/sites/45/2024/03/ipece_informe_239_06Mar2024.pdf. Acesso em: 01. mai. 2024.
- CEARÁ (Estado). Instituto de Pesquisa e Estratégia Econômica do Ceará – IPECE. **Enfoque econômico Nº 251 – Comportamento das Principais Receitas do Governo do Ceará no**

primeiro Bimestre de 2023. Disponível em: https://www.ipece.ce.gov.br/wp-content/uploads/sites/45/2023/04/Enfoque_Economico_N251_280423.pdf. Acesso em: 01. mai. 2024.

CLEMENTE, A.; CLEMENTE, L. T. Aplicação da metodologia Box-Jenkins para previsão do ICMS do estado do Paraná de agosto de 2011 a julho de 2012. **Economia & Tecnologia**, v. 27, p. 47-58, 2011. Disponível em: <https://revistas.ufpr.br/ret/article/viewArticle/25910> . Acesso em: 30 set. 2023.

FERREIRA, Roberto Tatiwa. **Modelos de análise de séries temporais para previsão do ICMS mensal do estado do Ceará.** Dissertação de Mestrado – Faculdade de Economia, Administração, Atuária e Contabilidade, Universidade Federal do Ceará, Fortaleza, 1996.

FERREIRA, Roberto Tatiwa. **Forecasting quarterly brazilian GDP growth rate with linear and nonlinear diffusion index models.** Tese (Doutorado em Economia) – Faculdade de Economia, Administração, Atuária e Contabilidade, Universidade Federal do Ceará, Fortaleza, 2005.

FERREIRA, Roberto Tatiwa. Ciclos econômicos na taxa de crescimento do ICMS. **Economia Aplicada** (São Paulo), v. 13, n. 1, p. 29-44, 2009. Disponível em: <https://www.scielo.br/j/ecoa/a/Njk94d489v3zPpJnZX7GBmx/>. Acessado em: 10. set. 2023.

FIGUEIREDO, Francisco Marcos R.; GUILLÉN, Osmani Teixeira de C. Forecasting brazilian consumer inflation with FAVAR models using targeted variables. Disponível em: <https://www.bcb.gov.br/secre/apres/FAVAR%20paper%20Figueiredo%20&%20Guillen%20prelim.pdf> . Acesso em: 10 jun. 2024.

MADHOU, Ashwin; SEWAK , Tayushma; MOOSA , Imad ; RAMIAH, Vikash. Forecasting the GDP of a small open developing economy: an application of FAVAR models. **Applied Economics**, v. 52, [New York]: Routledge, 2020. Disponível em: <https://doi.org/10.1080/00036846.2019.1679346> . Acesso em: 10 jun. 2024.

MATOS, Daniel A. S.; RODRIGUES, Erica C. **Análise fatorial.** Brasília: Enap, 2019. Disponível em: <https://repositorio.enap.gov.br/bitstream/1/4790/1/Livro%20An%c3%a1lise%20Fatorial.pdf> . Acessado em: 9 fev. 2024.

MENDONÇA, M. J. C. de, SACHSIDA, A., & MEDRANO, L. A. T. (2013). Um modelo econométrico para previsão de impostos no Brasil. **Economia Aplicada**, v. 17, n. 2, p. 295 – 329. Disponível em: <https://www.revistas.usp.br/ecoa/article/view/78254>. Acessado em: 26. set. 2023.

MENDONÇA, M. J. C. NASCIMENTO, I. F. do, & MEDRANO, L. A. T. (2022). **Um modelo para projeção de arrecadação tributária dos entes subnacionais.** Rio de Janeiro: Ipea, 2022. (Texto para Discussão, n. 2782). Disponível em: https://repositorio.ipea.gov.br/bitstream/11058/11239/1/td_2782.pdf . Acessado em: 26. set. 2023.

MOÇO, João Ricardo Cruz. **Estimação de um modelo de previsão do ICMS do Estado do Rio de Janeiro.** 2017. Dissertação (mestrado). Instituto de Economia – Universidade Federal do Rio de Janeiro, 2017. Disponível em:

<https://www.ie.ufrj.br/images/IE/PPGE/disserta%C3%A7%C3%B5es/2017/Jo%C3%A3o%20Ricardo%20Cruz%20Mo%C3%A7o.pdf> . Acesso em: 17 ago. 2023.

MORETTIN, Pedro Alberto e TOLOI, Clélia Maria de Castro. **Análise de séries temporais**. São Paulo: Blucher, 2022

PACCAGNINI, Alessia. Forecasting with FAVAR: macroeconomic versus financial factors. **NBP Working Paper**, n. 256, 2017.

PECEGUINI, E.E. **Análise Comparativa de Métodos de Previsão Aplicados à Arrecadação do ICMS – Estado de São Paulo**. (Aperfeiçoamento/Especialização em Economia do Setor Público – Fundação Instituto de Pesquisas Econômicas). São Paulo. 2001. Disponível em: http://www.planejamentotributario.ufc.br/artigo_ORCAMENTOS_SISTEMAS_DE_INFORMACAO_SOBRE_ADMINISTRACAO_FINANCEIRA_PUBLICA.pdf . Acessa em: 23 set. 2023.

PESSOA, Filipe de Moraes Cangussu; CORONEL, Daniel Arruda; DE LIMA, João Eustáquio. Previsão de arrecadação de ICMS para o estado de Minas Gerais: uma comparação entre modelos Arima e Arfima. **Revista Brasileira de Gestão e Desenvolvimento Regional**, v. 9, n. 2, 2013.

MORAIS, Hugo Azevedo Rangel de et al. O reflexo do Auxílio Emergencial nas unidades federativas brasileiras como manutenção das receitas tributárias no período pandêmico. **GeSec: Revista de Gestao e Secretariado**, v. 14, n. 3, 2023.

SOUSA, K. R., CORREIA, J. J. A., SANTOS, R. R. dos, & SANTOS, J. F. dos. (2019). Previsão de receitas tributárias: uma análise do ICMS no Ceará por meio de séries temporais. **Refas – Revista Fatec Zona Sul**, (5), 1–17. Disponível em: <https://revistarefas.com.br/RevFATECZS/article/view/323>. Acesso em: 25 set. 2023.

STOCK, James H.; WATSON, Mark W. Forecasting using principal components from a large number of predictors. **Journal of the American statistical association**, v. 97, n. 460, p. 1167-1179, 2002.

Stock, J. H. & Watson, M. W. Diffusion indexes. **Technical working paper 6702**. National Bureau of Economic Research, 1998. Disponível: <https://ideas.repec.org/p/nbr/nberwo/6702.html> . Acessado em: 11. fev. 2024.

Stock, J. H. & Watson, M.W. Macroeconomic forecasting using diffusion indexes. **Journal of Business & Economic Statistics**, 20, 2, 147–162, 2002. Disponível em: https://scholar.harvard.edu/files/stock/files/macroeconomic_forecasting_using_diffusion_indexes.pdf . Acessado em: 10. fev. 2024.

VIEIRA, Fausto; CHAGUE, Fernando; FERNANDES, Marcelo. Forecasting the brazilian yield curve using forward-looking variables. **International Journal of Forecasting**. v. 33, p. 121-131, 2017.