

De Mendonça, Mário Jorge Cardoso; Do Nascimento, Igor Ferreira; Medrano, Luis Alberto Toscano

**Working Paper**

## Um modelo para projeção de arrecadação tributária dos entes subnacionais

Texto para Discussão, No. 2782

**Provided in Cooperation with:**

Institute of Applied Economic Research (ipea), Brasília

*Suggested Citation:* De Mendonça, Mário Jorge Cardoso; Do Nascimento, Igor Ferreira; Medrano, Luis Alberto Toscano (2022) : Um modelo para projeção de arrecadação tributária dos entes subnacionais, Texto para Discussão, No. 2782, Instituto de Pesquisa Econômica Aplicada (IPEA), Brasília,  
<https://doi.org/10.38116/td2782>

This Version is available at:

<https://hdl.handle.net/10419/265301>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

TEXTO PARA DISCUSSÃO

2782

UM MODELO PARA PROJEÇÃO DE  
ARRECADAÇÃO TRIBUTÁRIA DOS  
ENTES SUBNACIONAIS

MARIO JORGE MENDONÇA  
IGOR FERREIRA DO NASCIMENTO  
LUIS ALBERTO MEDRANO



**UM MODELO PARA PROJEÇÃO DE  
ARRECAÇÃO TRIBUTÁRIA DOS ENTES  
SUBNACIONAIS**

**MARIO JORGE MENDONÇA<sup>1</sup>  
IGOR FERREIRA DO NASCIMENTO<sup>2</sup>  
LUIS ALBERTO MEDRANO<sup>3</sup>**

---

1. Técnico de planejamento e pesquisa na Diretoria de Estudos e Políticas Regionais, Urbanas e Ambientais do Instituto de Pesquisa Econômica Aplicada (Dirur/Ipea). *E-mail*: <mario.mendonca@ipea.gov.br>.

2. Assistente de pesquisa na Dirur/Ipea; e professor do Instituto Federal do Piauí. *E-mail*: <igor.ferreira.n@gmail.com>.

3. Professor adjunto do Departamento de Matemática da Universidade Federal Rural do Rio de Janeiro (UFRRJ). *E-mail*: <lmedrano@ufrj.br>.

**Governo Federal**

**Ministério da Economia**

**Ministro** Paulo Guedes

**ipea** Instituto de Pesquisa  
Econômica Aplicada

Fundação pública vinculada ao Ministério da Economia, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiros – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

**Presidente**

**ERIK ALENCAR DE FIGUEIREDO**

**Diretor de Desenvolvimento Institucional**

**ANDRÉ SAMPAIO ZUVANOV**

**Diretor de Estudos e Políticas do Estado, das  
Instituições e da Democracia**

**FLÁVIO LYRIO CARNEIRO**

**Diretor de Estudos e Políticas Macroeconômicas**  
**MARCO ANTÔNIO FREITAS DE HOLLANDA CAVALCANTI**

**Diretor de Estudos e Políticas Regionais,  
Urbanas e Ambientais**

**NILO LUIZ SACCARO JUNIOR**

**Diretor de Estudos e Políticas Setoriais de  
Inovação e Infraestrutura**

**JOÃO MARIA DE OLIVEIRA**

**Diretor de Estudos e Políticas Sociais**

**HERTON ELLERY ARAÚJO**

**Diretor de Estudos e Relações Econômicas e  
Políticas Internacionais**

**PAULO DE ANDRADE JACINTO**

**Assessor-chefe de Imprensa e  
Comunicação (substituto)**

**JOÃO CLÁUDIO GARCIA RODRIGUES LIMA**

**Ouvidoria:** <http://www.ipea.gov.br/ouvidoria>

**URL:** <http://www.ipea.gov.br>

## Texto para Discussão

Publicação seriada que divulga resultados de estudos e pesquisas em desenvolvimento pelo Ipea com o objetivo de fomentar o debate e oferecer subsídios à formulação e avaliação de políticas públicas.

© Instituto de Pesquisa Econômica Aplicada – **ipea** 2022

Texto para discussão / Instituto de Pesquisa Econômica  
Aplicada.- Brasília : Rio de Janeiro : Ipea , 1990-

ISSN 1415-4765

1. Brasil. 2. Aspectos Econômicos. 3. Aspectos Sociais.  
I. Instituto de Pesquisa Econômica Aplicada.

CDD 330.908

As publicações do Ipea estão disponíveis para *download* gratuito nos formatos PDF (todas) e EPUB (livros e periódicos).  
Acesse: <http://www.ipea.gov.br/portal/publicacoes>

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou do Ministério da Economia.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

JEL: C32; C38; C51; C52; C53.

DOI: <http://dx.doi.org/10.38116/td2782>

# SUMÁRIO

SINOPSE

ABSTRACT

1 INTRODUÇÃO .....	6
2 REFERENCIAL TEÓRICO.....	11
3 METODOLOGIA .....	17
4 TRATAMENTO DA BASE DE DADOS E ESTRATÉGIA DE MODELAGEM.....	24
5 CONSIDERAÇÕES FINAIS E IMPLICAÇÕES DE POLÍTICAS .....	40
REFERÊNCIAS .....	42
APÊNDICE A .....	46
APÊNDICE B .....	47
APÊNDICE C .....	48

## SINOPSE

A previsão de receitas governamentais é de extrema relevância para uma adequada execução orçamentária, pois uma boa acurácia na estimação permite que seja estipulado um nível de despesa que atenda as demandas da população. Utilizando-se os dados divulgados pelo Conselho Nacional de Política Fazendária (Confaz), este estudo tem por objetivo geral realizar projeções mensais da arrecadação tributária estadual desagregada com o uso do modelo fatorial dinâmico. Essas projeções são condicionadas aos cenários macroeconômicos oficiais de médio prazo relacionados a preços e atividade econômica. Os resultados obtidos foram razoáveis, ficando o erro percentual acumulado para seis meses em torno de 10% – considerando-se amostra composta de mais de quatrocentas séries contendo padrões de tendência e sazonalidade bastante dispare, além de problema de quebra estrutural, erro de medida, dados faltantes etc. Este estudo apresenta uma proposta de projeções de arrecadação tributária estadual que beneficiará os gestores públicos no planejamento fiscal.

**Palavras-chave:** arrecadação tributária estadual; modelo de espaço-estado; modelo fatorial dinâmico; previsão.

## ABSTRACT

The forecasting of government revenues is extremely important for an adequate budget execution, since a good accuracy in the estimation allows the stipulation of an expenditure level that meets the demands of the population. Using data released by the National Council of Treasury Policy, the general objective of this study is to make monthly projections of disaggregated state tax revenue using the dynamic factorial model. These projections are conditioned to the official medium-term macroeconomic scenarios related to prices and economic activity. The results obtained were reasonable, and the accumulated percentage error for six months was around 10%. This considering a sample composed of more than 400 series containing quite disparate trend and seasonality patterns, besides the problem of structural breaks, measurement error, missing data, etc. This study presents a proposal for state tax revenue projections that will benefit public managers in fiscal planning.

**Keywords:** state tax collection; state-space model; dynamic factor model; forecasting.

## 1 INTRODUÇÃO

### 1.1 Contextualização do tema

Em uma Federação, a estabilidade macroeconômica depende da situação fiscal não somente do governo federal, mas também dos governos subnacionais. Nesse contexto, questões importantes como a sustentabilidade da dívida pública e o desenho da política fiscal são conceitos que guardam relação estrita na medida em que o tamanho do endividamento limita a atuação da política fiscal, forçando-a a se tornar mais restritiva. Cabe ter em mente que o endividamento público é tema relevante desde 2015, quando o antigo Ministério da Fazenda buscou a renegociação das dívidas dos estados com a União, contemplando as exigências do Supremo Tribunal Federal, visando ao reequilíbrio das contas públicas, sem aumento de custos para o governo federal, sendo compatível com a volta da confiança e do crescimento econômico. Essa renegociação ocorreu simultaneamente às discussões sobre compromissos fiscais e contrapartidas que possam corrigir desequilíbrios estruturais nas finanças estaduais, e não apenas problemas conjunturais. Um elemento que deve dar suporte à política de renegociação da dívida dos estados diz respeito ao conhecimento acerca da capacidade de arrecadação de tributos a longo prazo desses entes subnacionais, visto que essa capacidade é o fator preponderante capaz de fazer frente ao endividamento. Portanto, qualquer que seja a proposta de renegociação da dívida, deve-se considerar um horizonte longo acerca da trajetória da arrecadação. A obtenção de evidências empíricas quanto à capacidade arrecadatória dos entes subnacionais possibilitará aos gestores públicos a adoção de medidas que visem a um melhor gerenciamento do endividamento público.

Desde que se iniciaram as primeiras análises feitas para estimar a elasticidade da carga tributária bruta, os estudos evoluíram em direção ao tratamento das séries individuais de tributos, aumentando assim a capacidade de controle para o gestor de política. Exemplo disso é o estudo elaborado por Mendonça, Sachsida e Medrano (2013), cuja amostra de tributos contempla algumas das séries do imposto de renda (IR) e do imposto sobre produtos industrializados (IPI), receita previdenciária, além da série agregada do Imposto sobre a Circulação de Mercadorias e Prestação de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação (ICMS). Mais recentemente, Gadelha, Lima e Polli (2020) aplicaram o modelo de combinação de previsão (Bates e Granger, 1969) em uma base de dados de arrecadação bruta de tributos federais administrados pela Secretaria Especial da Receita Federal do Brasil, obtendo-se projeções com grande grau de acurácia. No entanto, cabe notar que boa parte dessas séries de tributos federais é bem-comportada do ponto de vista estatístico, apresentando, em muitos casos, tendência e sazonalidade bem definidas.

Infelizmente, essa não é a situação comum com que se depara o gestor de política, de modo que a necessidade de extrair informação vai bem além daquela contida na arrecadação agregada ou em séries de tributos com formato bem definido. Este é o caso dos tributos de competência dos entes subnacionais como os estados e o Distrito Federal. Mais especificamente, faz-se referência às séries de tributos dos entes subnacionais administradas pelo Conselho Nacional de Política Fazendária (Confaz). Esse conjunto de dados consta de dezessete séries genuínas de tributos para cada uma das 27 Unidades da Federação (UFs). Cabe ter em mente que as dinâmicas dessas séries podem apresentar entre elas comportamentos bem distintos. Por exemplo, a série do ICMS do mercado atacadista tem uma dinâmica bem distinta da série do ICMS do setor de transporte. Existe forte heterogeneidade nesses entes subnacionais no que diz respeito às suas economias, o que se reflete também na dinâmica de arrecadação. Além disso, um estado que tenha sua economia baseada no setor primário pode ter, no ICMS desse setor, características distintas daquelas observadas para a mesma categoria de ICMS de outro estado no qual esse setor tem pouca relevância.

Além disso, deve-se considerar que a arrecadação de um ente subnacional, embora tenha como fator principal sua capacidade econômica, também deriva da eficiência de arrecadação deste ente. Portanto, quando se visa ao tratamento dos tributos dos entes subnacionais, é necessário que se proponham instrumentos mais flexíveis, capazes de lidar com séries de tempo que apresentam dinâmicas bastante complexas e díspares.

## 1.2 Problema de pesquisa

Este estudo apresenta o seguinte problema de pesquisa: as técnicas bayesianas podem fornecer projeções de arrecadação tributária estadual mais acuradas, contribuindo assim para o planejamento orçamentário e financeiro dos estados e do Distrito Federal?

## 1.3 Hipótese da pesquisa

De modo geral, uma boa previsão precisa: i) ter acurácia, ou seja, o erro de previsão *ex-post* precisa ser o menor possível; ii) enviesada, isto é, a previsão não deve ser nem otimista nem pessimista; e iii) racional, no sentido de que toda a informação disponível é utilizada na previsão.

Em vista disso, a hipótese que conduz esta pesquisa reside no fato de que o uso de técnicas bayesianas pode melhorar a acurácia das previsões da arrecadação tributária estadual, superando inclusive limitações já constatadas na literatura em relação ao uso de outros métodos para essa mesma finalidade, como modelagem Box-Jenkins ou modelo Holt-Winters aditivo. Além disso, as

técnicas bayesianas apresentam propriedades de ausência de viés e permitem o uso ótimo de toda a informação disponível para o exercício de previsão.

#### 1.4 Objetivos geral e específicos

O objetivo geral deste estudo é apresentar uma proposta metodológica visando à projeção da arrecadação das séries de tributos dos entes subnacionais (estados e Distrito Federal).

Com intuito de contemplar todos esses fatores, os quais incorporam tendência, sazonalidade e diversidade de fatos geradores, como objetivo específico propõe-se nesse estudo uma metodologia para projeção de arrecadação de tributos dos entes subnacionais com base no emprego do modelo fatorial dinâmico (MFD). Uma razão fundamental para o uso do MFD é que este permite obter ganho de informação importante por meio da redução de dimensionalidade (isto é, de aumento no grau de agregação) das variáveis. Nesse caso, a informação contida nas inter-relações entre as variáveis é “passada para um conjunto menor de variáveis” composto pelos fatores, os quais possuem a habilidade de identificar a “verdadeira” estrutura dos dados. Do ponto de vista prático, como é um modelo multivariado, o MFD tem a vantagem de projetar todas as séries simultaneamente e é estimado aqui com base na abordagem bayesiana (Gamerman e Lopes, 2006; Kim e Nelson, 1999; West e Harrison, 1997; Durbin e Koopman, 2001), cujas vantagens serão externadas à frente.

Cabe mencionar que a especificação do modelo fatorial dinâmico se diferencia por sua capacidade de lidar com diversos problemas relacionados à estrutura das séries, como o tratamento da sazonalidade, emprego de coeficientes variáveis e emprego de tratamento específico para lidar com a quebra estrutural devido à pandemia da covid-19 (coronavírus). No método adotado neste estudo, a sazonalidade é vista como um conjunto de parâmetros a mais dentro do modelo. Tal procedimento permite obter estimativas significativamente mais ajustadas aos dados, assim como projeções mais confiáveis, sobretudo nas séries de tributos dominadas por elementos sazonais. A introdução de coeficientes variáveis se mostra plausível *a priori*, tendo em vista que as sucessivas mudanças existentes na tributação podem impor mudanças aos coeficientes mesmo que não tão abruptas no curto prazo, mas que no médio e longo prazo geram efeitos importantes. Por fim, adota-se um tratamento específico para lidar com o efeito da pandemia da covid-19 sobre a arrecadação tributária.

#### 1.5 Delimitação do escopo do estudo

Por estar voltado à elaboração de previsões de arrecadação tributária estadual, este estudo possui algumas delimitações. A primeira advém do fato de que as receitas divulgadas no boletim

do Confaz serão o foco dos exercícios de previsões, ou seja, as projeções serão feitas apenas para a arrecadação das receitas tributárias estaduais, não se considerando aqui tributos federais ou municipais. A segunda delimitação reside no fato de que se pretende fornecer aos formuladores de políticas uma metodologia robusta de previsões de arrecadação de tributos estaduais considerando-se um horizonte de médio prazo. A terceira é de natureza metodológica, pois as previsões de arrecadação de tributos estaduais a serem aqui tratadas estão condicionadas aos cenários de nível de atividade econômica e de índices de preços. Esses cenários são fundamentais, uma vez que a evolução das bases tributárias ou os custos/rendimentos de novas medidas de receita são geralmente previstos usando indicadores macroeconômicos bem escolhidos. Além disso, quando a atividade econômica é mais dinâmica, existem mais atividades a serem tributadas.

## 1.6 Justificativa do tema

A previsão de receita, a qual se refere ao uso de técnicas analíticas para produzir estimativas de fluxos financeiros futuros para o governo, é de importância crucial para os gestores públicos, pois ajuda na formulação de políticas públicas sólidas que dão suporte à sustentabilidade fiscal, bem como no planejamento da programação financeira de despesas com base nas receitas projetadas. Assim, o exercício da previsão das receitas governamentais é de extrema relevância para uma adequada execução orçamentária, pois uma boa acurácia na estimação permite que seja estipulado um nível de despesa que atenda as demandas da população. Essa importância é explicitada no artigo 11 da Lei Complementar nº 101, de 4 de maio de 2000 (Lei de Responsabilidade Fiscal – LRF), a qual obriga os entes da Federação a realizarem a previsão de arrecadação: “Art. 11. Constituem requisitos essenciais da responsabilidade na gestão fiscal a instituição, previsão e efetiva arrecadação de todos os tributos da competência constitucional do ente da Federação” (Brasil, 2000).

Além disso, a LRF preceitua que as projeções de receitas devem ser divulgadas por todos os entes da Federação por meio da Lei de Diretrizes Orçamentárias (LDO) e da Lei Orçamentária Anual (LOA), conforme estabelecido no §1º do artigo 4º e no inciso I do artigo 5º.

Art. 4º A lei de diretrizes orçamentárias atenderá o disposto no § 2º do art. 165 da Constituição e: (...)

§ 1º Integrará o projeto de lei de diretrizes orçamentárias Anexo de Metas Fiscais, em que serão estabelecidas metas anuais, em valores correntes e constantes, relativas a receitas, despesas, resultados nominal e primário e montante da dívida pública, para o exercício a que se referirem e para os dois seguintes. (...)

Art. 5º O projeto de lei orçamentária anual, elaborado de forma compatível com o plano plurianual, com a lei de diretrizes orçamentárias e com as normas desta Lei Complementar:

I - conterá, em anexo, demonstrativo da compatibilidade da programação dos orçamentos com os objetivos e metas constantes do documento de que trata o § 1º do art. 4º. (Brasil, 2000).

Pode-se, dessa forma, resumir que a LRF exige que todos os entes da Federação apresentem as previsões para as suas receitas no Projeto de Lei de Diretrizes Orçamentárias (PLDO) e no Projeto de Lei Orçamentária Anual (PLOA), sendo que o anexo da LOA atualiza os valores previstos no anexo de metas fiscais da LDO.

Conforme destacado no *Manual da Receita Nacional*, elaborado pela Secretaria do Tesouro Nacional (Brasil, 2008, p. 169-170), é importante destacar que a alocação de recursos orçamentários nas despesas públicas é função da previsão das receitas orçamentárias. Assim, deve-se atentar para o correto uso dos modelos de projeção de receitas, uma vez que isto será fundamental para a fixação das despesas na LOA e para a execução do orçamento.

Na fixação das despesas no PLOA, a dotação orçamentária depende da previsão da arrecadação. Dessa forma, a lei fixa os gastos orçamentários como função da receita prevista, podendo ocorrer três situações: i) superestimar as receitas; ii) estimar corretamente; e iii) subestimar as receitas. Se a receita for superestimada, a despesa também o será, sendo necessário um contingenciamento de dotações. Similarmente, se a receita for subestimada, a despesa também o será, podendo haver solicitação de créditos adicionais por excesso de arrecadação. No ano seguinte, se não forem utilizados os recursos arrecadados a maior, poderá haver pedido de crédito adicional por superávit financeiro, o que dependerá de aprovação parlamentar.

Uma correta projeção de receitas também é fundamental para subsidiar a elaboração da programação financeira e do cronograma mensal de desembolso, previsto no art. 8º da LRF: "Até trinta dias após a publicação dos orçamentos, nos termos em que dispuser a lei de diretrizes orçamentárias e observado o disposto na alínea c do inciso I do art. 4º, o Poder Executivo estabelecerá a programação financeira e o cronograma de execução mensal de desembolso" (Brasil, 2000).

Dessa forma, a projeção inadequada da receita poderá ocasionar o1 contingenciamento de créditos com a consequente limitação de empenhos, conforme prevê o art. 9º da LRF:

se verificado, ao final de um bimestre, que a realização da receita poderá não comportar o cumprimento das metas de resultado primário ou nominal estabelecidas no Anexo de Metas Fiscais, os Poderes e o Ministério Público promoverão, por ato próprio e nos

montantes necessários, nos trinta dias subsequentes, limitação de empenho e movimentação financeira, segundo os critérios fixados pela lei de diretrizes orçamentárias (Brasil, 2000).

Havendo reestimativa de receitas com base no excesso de arrecadação e na observação da tendência do exercício, pode-se solicitar crédito adicional. Supondo que essa tendência não se realize e que o Poder Legislativo do ente tenha aprovado o crédito adicional, torna-se possível a emissão de empenho sem a correspondente fonte de recursos, que pode gerar restos a pagar para o exercício seguinte ou cancelamento do empenho. Por outro lado, se houver reestimativa de receita prevendo uma frustração na arrecadação que não se concretize, poderá ocorrer uma alteração no planejamento governamental, e limitação de empenhos, desnecessária.

Portanto, previsões confiáveis de receitas são insumos relevantes para o processo de elaboração do orçamento para cada ente da Federação. E embora os governos obtenham receitas de várias fontes, os impostos normalmente são o principal mecanismo de financiamento.

## 1.7 Organização do estudo

Além desta introdução, este Texto para Discussão (TD) está estruturado da seguinte forma. A segunda seção apresenta referencial teórico relacionado às projeções de tributos estaduais, enquanto a terceira descreve os diferentes aspectos metodológicos que compõem o modelo fatorial dinâmico, destacando-se as etapas de estimação e projeção. A quarta seção descreve a base de dados e as estratégias que podem ser usadas para avaliar a melhor maneira de desagregar o total das séries de tributos em subconjuntos, ao mesmo tempo que também perfaz todo um procedimento estatístico para avaliar qual das opções de desagregação produz o melhor resultado em termos de previsão. A seguir, tendo por base a melhor estratégia de desagregação, na quinta seção são apresentados os resultados da projeção das séries de tributos estaduais de 2021 até dezembro de 2025. As considerações finais, englobando as implicações de política e os futuros desdobramentos desta pesquisa, são postas na última seção.

## 2 REFERENCIAL TEÓRICO

O referencial teórico encontra-se dividido em duas subseções. A primeira subseção trata de analisar alguns estudos voltados para a previsão da arrecadação de tributos estaduais no Brasil, enquanto a segunda destaca a contribuição deste estudo à literatura e ao processo orçamentário.

Apesar de já existirem estudos voltados para a arrecadação de tributos federais no Brasil (Melo, 2001; Siqueira, 2002; Campos, 2009; Benelli, 2013; Mendonça, Sachsida e Medrano, 2013; Gadelha, Lima e Polli, 2020), poucos são voltados para a previsão da arrecadação de tributos estaduais. E, conforme será comentado a seguir, esses estudos são direcionados para a realização da projeção da arrecadação agregada do ICMS; contudo, não de forma desagregada, como é feito aqui ao se utilizarem os dados desagregados de arrecadação tributária estadual divulgados pelo Confaz. Além disso, a maior parte desses estudos fizeram uso da modelagem Box-Jenkins, e poucos avançaram no uso de técnicas bayesianas para projeções de arrecadação tributária.

## 2.1 Estudos sobre previsões da arrecadação de tributos estaduais no Brasil com uso de modelos univariados

Peceguini (2001) comparou métodos de previsão da arrecadação de ICMS anual e mensal no estado de São Paulo para o ano de 2000, usando dados para o período de 1995 a 1999. Dentre os métodos univariados utilizados – decomposição clássica, linha de tendência e *autoregressive integrated moving average* (ARIMA) –, o método da decomposição clássica apresentou, de forma geral, melhores previsões em relação ao modelo base de comparação, denominado modelo aritmético.

Siqueira (2002) realiza uma comparação entre os modelos univariados e o método dos indicadores,<sup>1</sup> levando-se em conta um conjunto de nove tributos federais mais o ICMS. Os dados utilizados para as estimações dos modelos abrangeram o período de 1989 a 2000, sendo que o ano de 2001 foi usado como base de comparação dos resultados. Os resultados obtidos indicaram um melhor desempenho preditivo nos modelos ARIMA em oito das séries estudadas. Em duas delas – a Contribuição para o Financiamento da Seguridade Social (Cofins) adicionada ao Programa de Integração Social (PIS) e outras contribuições sociais – o método dos indicadores obteve melhores resultados. Entre as justificativas da superioridade do método convencional da Receita Federal do Brasil (RFB) em relação ao ARIMA nesses dois últimos casos, destacaram-se: i) presença de quebras estruturais nas séries; ii) alterações legais em variáveis significantes da arrecadação (alíquotas, bases de cálculo etc.); e iii) deficiência dos métodos quantitativos em capturar informações importantes que possam alterar o padrão histórico das séries.

---

1. Trata-se de um método utilizado pela Secretaria Especial da Receita Federal do Brasil como metodologia de projeção de receitas orçamentárias. Para obtenção de um resumo deste método, ver Gadelha, Lima e Polli (2020).

**TEXTO para DISCUSSÃO**

Guaragna e Mello (2002) estimaram a arrecadação do ICMS no estado do Rio Grande do Sul (RS) para os anos de 1999 a 2001, tomando-se como base os dados de 1989 a 1998. Os métodos univariados usados apresentaram resultados satisfatórios em relação aos valores observados.

Corvalão (2002) aplicou métodos multivariados – modelo de regressão dinâmica – para a previsão do ICMS de Santa Catarina. Usando dados como faturamento da indústria, consumo de energia elétrica e consultas ao serviço de proteção ao crediário (SPC) na comparação com o método oficial da Secretaria de Fazenda (Sefaz) de Santa Catarina, o qual se baseava em modelos univariados, foi possível reduzir os erros de previsão.

Com a metodologia Box-Jenkins, Camargos (2008) usou modelos da classe *seasonal autoregressive integrated moving average model* (SARIMA) para as séries dos tributos estaduais de São Paulo – ICMS, Imposto sobre a Propriedade de Veículos Automotores (IPVA), Imposto sobre a Transmissão Causa Mortis e Doação de Quaisquer Bens ou Direitos (ITCD) e Imposto de Transmissão de Bens Imóveis (ITBI) – e taxas.

Considerando-se a importância da arrecadação do ICMS para o orçamento do governo do estado do Paraná e a necessidade de se realizarem previsões para o planejamento das despesas, Clemente e Clemente (2011) procuraram prever a arrecadação desse tributo para o período de agosto de 2011 a julho de 2012. Aplicando-se a metodologia Box-Jenkins para modelar a série mensal de ICMS de janeiro de 2000 a julho de 2011, a qual foi diferenciada para se tornar estacionária, e, a partir da análise de resíduo e dos testes de significância dos parâmetros, elegeu-se a identificação SARIMA  $(2,1,0) \times (0,1,1)_{12}$  para previsão. Os resultados obtidos mostraram que a arrecadação do ICMS 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.

Pessoa, Coronel e Lima (2013) analisaram a capacidade dos modelos ARIMA e *autoregressive fractionally integrated moving average model* (ARFIMA) em prever a arrecadação do ICMS do estado de Minas Gerais. Não houve total predominância de um método sobre o outro, porém ambos os modelos tiveram bons resultados de previsão.

Chain *et al.* (2015) estimaram um modelo preditivo da arrecadação do ICMS pelo governo de Minas Gerais, abrangendo o período de janeiro de 1998 a agosto de 2011, utilizando-se a metodologia Box-Jenkins. Os resultados obtidos mostraram que o método *seasonal autoregressive integrated moving average with exogenous factors* (SARIMAX), ao considerar eventos diferentes do padrão histórico da série, apresentou melhor desempenho em relação às medidas de erro quando comparado aos métodos ARIMA, ARFIMA e SARIMA. Concluiu-se que a modelagem de valores atípicos contribuiu para uma melhor previsão das receitas de ICMS em Minas Gerais.

Duarte, Souza e Girão (2014) aplicaram o modelo de alisamento exponencial de Holt-Winters para prever a arrecadação mensal do ICMS do estado da Paraíba nos anos de 2013, 2014 e 2015 para comparar a previsão de 2013 com os valores reais arrecadados pelo estado. Os resultados obtidos mostraram que os valores arrecadados em 2013 são muito próximos com os previstos pelo modelo e que, nos períodos de finais de anos, ocorrem variações maiores, ou seja, há um aumento significativo na arrecadação do ICMS, em que supostamente acontece um maior consumo por parte da população devido, provavelmente, às festividades natalinas e férias escolares.

O estudo de Azevedo, Silva e Gatsios (2017) teve por objetivo verificar se a metodologia ARIMA prevê com mais exatidão as arrecadações de ICMS do que as metodologias utilizadas pelos entes federativos. Para tanto, a amostra foi composta dos seis estados brasileiros (SP, Minas Gerais, Rio de Janeiro, Rio Grande do Sul, Paraná e Bahia), os quais conjuntamente corresponderam a 69,7% do ICMS arrecadado pelos 27 estados. Assim sendo, foram estimados os valores das arrecadações para os anos de 2012 e 2013 a partir do horizonte temporal de 1995 a 2011. Os resultados apontaram que a utilização do ARIMA aumentou a acurácia da previsão de arrecadação do ICMS para todos os estados analisados. Ressalta-se, entretanto, que os modelos de séries temporais não são aconselháveis quando existem quebras estruturais de modo a alterar os aspectos macroeconômicos. Assim, caberiam às autoridades nos entes subnacionais a capacitação de seu quadro de funcionários para a utilização dessas metodologias de séries temporais de modo a complementar os procedimentos de estimação da arrecadação. Por fim, apesar das séries temporais terem melhorado a acurácia das estimativas, deve-se atentar para o fato de que a cada ano a série se altera, e novos testes devem ser realizados para a construção de novos modelos.

## **2.2 Estudos sobre previsões de arrecadação de tributos estaduais no Brasil com uso de modelos multivariados e modelos bayesianos**

Seguindo a mesma linha de pesquisa de Corvalão (2002), Castanho (2011) constatou que os modelos multivariados para projeções da arrecadação de ICMS apresentaram resultados mais acurados do que os métodos simples – como linha da tendência – que eram utilizados pela Sefaz do Espírito Santo. Ao se utilizarem modelos de alisamento exponencial de Holt-Winter e Box-Jenkins, os resultados obtidos mostraram que o aquele é mais eficiente que este na previsão de períodos de curto prazo em momentos de estabilidade econômica.

Em um estudo aplicado para o estado do Rio de Janeiro, Lima, Ferreira e Honaiser (2013) fizeram uso de técnicas bayesianas, dentre outros métodos, para a previsão de arrecadação tributária estadual de diversos tributos, dentre os quais, ICMS e IPVA. Os modelos univariados

utilizados foram: SARIMA e modelos estruturais. Por sua vez, os modelos multivariados utilizados foram: modelo vetorial autorregressivo bayesiano (*bayesian vector autoregressive model* – BVAR) e modelo vetorial autorregressivo bayesiano com mecanismo de correção de erros (*bayesian vector error correction model* – BVEC). Em particular, para a série de ICMS, os modelos multivariados utilizados levaram em consideração as seguintes covariáveis para o estado do Rio de Janeiro: população, produção industrial, comércio varejista e produção de petróleo, além do consumo de energia elétrica da região Sudeste.

Moço (2017) comparou a acurácia preditiva do modelo oficial de previsão de ICMS da Sefaz do estado do Rio de Janeiro. A motivação principal do estudo foi a piora dos resultados de previsão do modelo oficial nos anos de 2015 e 2016 (erros anuais de, respectivamente, 14,75% e 12,64%). Em face disso, o estudo teve por objetivo estimar modelos econométricos univariados, multivariados e com métodos de combinação, bem como comparar seus resultados com o modelo oficial utilizado pela Sefaz do Rio de Janeiro. As estimações foram feitas usando dados de janeiro de 2003 até junho de 2015 e as previsões realizadas para julho de 2015 até dezembro de 2016, tendo as estatísticas de erro de previsão computadas apenas com as previsões para os meses de 2016. Os resultados apontaram para relevante aumento da acurácia preditiva dos modelos estimados quando comparados com o oficial. Em especial, foi destacado que o melhor modelo de combinação de previsão obtido gerou os resultados mais acurados para quatro das cinco estatísticas de erro consideradas, enquanto que para a estatística restante, relacionada ao erro percentual anual, o modelo multivariado foi o mais acurado.

### 2.3 Contribuições

Este estudo apresenta, a princípio, cinco contribuições relevantes. Primeiro – assim como em Lima, Ferreira e Honaiser (2013) –, esse estudo contribui ao propor o uso de técnicas de inferência bayesiana nas projeções de arrecadação tributária estadual. Essas técnicas possuem um reconhecimento sedimentado na literatura quanto ao seu potencial metodológico para previsões quando comparado com outras metodologias clássicas, além de superarem problemas relacionados a estacionariedade e presença de quebras estruturais nas séries de dados. Contudo, enquanto Lima, Ferreira e Honaiser (2013) utilizam abordagens bayesianas multivariadas (BVAR e BVEC), neste estudo usa-se o modelo fatorial dinâmico para as projeções da arrecadação tributária estadual.

Segundo, neste estudo são apresentados resultados importantes acerca dos ganhos de eficiência e de acurácia ao se utilizar uma mesma abordagem metodológica bayesiana de previsão de receitas ao longo do processo orçamentário dos estados e Distrito Federal. E, ao contrário de estudos anteriores, que para cada ente da Federação foram utilizadas metodologias distintas

de previsão, este possibilita aos gestores públicos dos entes subnacionais o uso de uma mesma metodologia de previsão em código livre (*software R*) para a previsão da arrecadação tributária estadual. Esse aspecto é importante pois possibilita uma capacitação mais adequada dos servidores públicos estaduais na elaboração dos exercícios de projeções de arrecadação tributária, sendo uma preocupação considerada em Azevedo, Silva e Gatsios (2017).

Terceiro, sob o aspecto de formulação de políticas, esse estudo é inovador para o processo orçamentário dos entes subnacionais pois, ao contrário de estudos anteriores que se restringiram a realizarem previsões da arrecadação agregada de ICMS, aqui realizam-se as previsões de dezessete séries temporais de tributos estaduais para cada um dos estados e Distrito Federal, usando-se uma mesma metodologia de previsão e tendo o boletim do Confaz<sup>2</sup> como fonte de dados primários e públicos para essas previsões.

Quarto, este estudo é uma importante referência metodológica desenvolvida em conjunto pelas equipes do Instituto de Pesquisa Econômica Aplicada (Ipea) e da Secretaria de Política Econômica (SPE), pois fornece previsões robustas da arrecadação tributária estadual para análises e tomadas de decisões no âmbito da Secretaria Especial da Receita Especial do Brasil e da Secretaria Especial do Tesouro e Orçamento, ambas do Ministério da Economia. Por exemplo, nesta última secretaria especial, a metodologia e as previsões da arrecadação tributária estadual apresentadas neste estudo se tornam insumos importantes para subsidiar os trabalhos das equipes da coordenação-geral das relações e análise financeira dos estados e municípios da Secretaria do Tesouro Nacional no que se refere à coordenação, supervisão e avaliação da execução das atividades relativas ao Programa de Reestruturação Fiscal, ao Programa de Acompanhamento Fiscal, à capacidade de pagamento e aos estudos fiscais, além do aperfeiçoamento do relacionamento financeiro da União com os entes subnacionais. E, como em uma Federação a estabilidade macroeconômica depende da situação fiscal dos entes subnacionais, os resultados deste estudo são também importantes para os trabalhos desenvolvidos pelas equipes da Secretaria de Orçamento Federal na elaboração das peças orçamentárias ao longo do ciclo orçamentário. Já na Secretaria Especial da Receita Federal do Brasil, por exemplo, a equipe do Centro de Estudos Tributários e Aduaneiros (Cetad) pode se beneficiar das previsões fornecidas segundo a metodologia aqui utilizada para o acompanhamento e avaliação do desempenho da arrecadação tributária, além da formulação e análise de propostas de política tributária e aduaneira.

Por fim, poucos estudos propuseram realizar previsões de arrecadação tributária estadual considerando um horizonte de médio prazo. A maior parte dos estudos consultados no referencial

---

2. Disponível em: <<https://bit.ly/3OGLgnl>>.

teórico objetivam estimar a arrecadação do ICMS no curto prazo. Não foram encontrados estudos similares à proposta aqui apresentada para estimar outras fontes de arrecadação tributária estadual. Além disso, as projeções de arrecadação tributária estadual a serem aqui reportadas encontram-se condicionadas aos cenários macroeconômicos de médio prazo.

### 3 METODOLOGIA

Justifica-se o desenvolvimento de metodologias a serem utilizadas na estimação de receitas públicas, em qualquer ente da Federação (União, estados, Distrito Federal e municípios), uma vez que a previsão de receitas determinará a capacidade de gastos desses entes, servindo de base para o valor de despesas a ser alocado em diferentes áreas (como educação, saúde e segurança pública) durante o processo orçamentário. É neste momento de elaboração das peças orçamentárias – LDO e LOA – que os formuladores de política (Poder Executivo) e os legisladores (Poder Legislativo) contribuem ao planejamento e ao controle social, ao analisarem aspectos relacionados à superestimação ou subestimação de receitas, visando ao equilíbrio orçamentário (Rubin, 2009).

Para serem confiáveis, as previsões de receita precisam resultar a partir da aplicação de processos e metodologias predefinidos. Essas previsões não devem ser sistematicamente otimistas ou pessimistas. Por um lado, as previsões de receita excessivamente otimistas podem forçar os governos a realizar cortes orçamentários não planejados e não coordenados ao longo do ano fiscal. Por outro lado, as previsões de receita que são excessivamente pessimistas podem restringir desnecessariamente os gastos públicos cruciais nas áreas de educação, saúde ou infraestrutura, por exemplo, além de poder levar o governo a tomar mais empréstimos do que deveria, incorrendo assim em pagamentos de juros da dívida pública.

Segundo Oliveira (2015, p. 804),

a escolha do modelo depende do comportamento temporal da arrecadação e da identificação de fatores que explicam razoavelmente alterações de um ano para outro, como inflação, crescimento econômico e outros fatores específicos. Em regra, o comportamento da arrecadação varia ao longo do ano e de receita para receita.

Dentro do escopo das metodologias consagradas na literatura para estimação de receitas, destaca-se o modelo fatorial dinâmico, o qual pertence à classe de modelos de variáveis latentes, que são ferramentas importantes para a análise de dados multivariados. Uma razão dessa importância é a redução de dimensionalidade que proporcionam. O objetivo da análise fatorial é reduzir as dimensões de modo que sejam apenas usadas as fontes de informação mais importantes.

A ideia é que a informação contida nas variáveis que são mais correlacionadas pode ser mapeada por um subconjunto menor de variáveis latentes denominado fator. Os fatores devem ser não correlacionados, pois representam fontes independentes de informação. A redução de dimensionalidade aumenta a eficiência computacional resultante dessa ausência de correlação. Dado que os fatores são ortogonais, sua matriz de covariância não condicional é diagonal.

Além disso, essa redução de dimensionalidade faz com que os modelos fatoriais sejam parcimoniosos quando comparados com outros que analisam dados multivariados, sem proporcionar essa redução de dimensionalidade, como os modelos vetoriais autorregressivos VAR (Hamilton, 1994, p. 291-350), que são muito usados na literatura econométrica e lidam melhor quando o conjunto de variáveis é restrito. Mais detalhes a respeito das vantagens do uso de modelos fatoriais podem ser encontrados em Bartholomew (1987).

Conforme é conhecido, o MFD é capaz de lidar com um vasto conjunto de séries simultaneamente. Essa é uma característica do modelo que é útil àquele pesquisador ou gestor que trabalha com grande quantidade de dados, e que tem necessidade de tomar decisão de forma rápida. Também pode ser de boa valia quando se tem pouca informação com respeito ao fato gerador, haja vista o modelo fatorial fazer uso eficiente das correlações existentes entre as variáveis.

### 3.1 Modelo de espaço de estudo

O modelo fatorial dinâmico utilizado neste trabalho é um caso particular dentro de uma abordagem mais geral denominada de modelo de espaço de estado (Petris, Petrone e Campagnoli, 2009). Tais modelos têm como características básicas a linearidade e normalidade para o distúrbio. No modelo de espaço de estado, a estimação e previsão podem ser obtidas recursivamente pelo que se conhece como filtro de Kalman. Os modelos de espaço de estado podem ser usados para modelar séries de tempo na presença de não estacionariedade, de mudanças estruturais e de padrões irregulares. Se a série se mostra bastante previsível, uma vez que suas características básicas se repetem regularmente, temos o caso de uma tendência e uma sazonalidade bastante regular com uma variabilidade ligeiramente crescente, o que facilita sua previsão futura.

O experimento se torna complexo quando as séries temporais não assumem um padrão regular e não apresentam propriedades de estacionariedade, mas pode incluir pontos de mudança na tendência ou quebras estruturais. A análise de séries de tempo não estacionárias com modelos *autoregressive moving-average* (ARMA) requer, pelo menos, uma transformação preliminar dos dados, de modo que eles adquiram aparência mais tratável. No entanto, é desejável ter meios para analisar os dados mais diretamente sem desconsiderar a instabilidade na média e na variância, as

## TEXTO para DISCUSSÃO

quebras estruturais e os saltos. Os modelos de espaço de estado incluem os modelos ARMA como um caso especial, mas este também pode ser aplicado em séries não estacionárias sem necessidade de uma transformação preliminar dos dados.

A estrutura mais geral dos modelos gaussianos lineares de espaço de estado nos quais estão incluídos modelos estruturais propostos neste estudo é especificada por uma distribuição *priori* normal para o vetor  $\theta$  de dimensão  $p$  no tempo  $t=0$ , de modo que  $\theta(0) \sim N_p(m_\theta, C_\theta)$  juntamente com o par de equações em que para cada  $t \geq 1$ , tem-se:

$$Y(t) = F(t) \theta(t) + v(t), \quad (1)$$

$$\theta(t) = G(t) \theta(t-1) + w(t). \quad (2)$$

Os distúrbios  $v(t)$  e  $w(t)$  e são independentes e gaussianos de modo que:

$$v(t) \sim N_m(0, V(t)),$$

$$w(t) \sim N_p(0, W(t)).$$

As matrizes  $G(t)$  e  $F(t)$  e são de ordem, respectivamente,  $p \times p$  e  $m \times p$ .  $v(0) \sim N_p(m_\theta, C_\theta)$   $\theta(0) \sim N_p(m_\theta, C_\theta)$ . A equação (1) é conhecida como equação de observação enquanto a equação (2), equação do estado.

Diferente do modelo definido pelas equações (1) e (2), o modelo de espaço de estado mais geral pode ser especificado a partir de uma distribuição *priori* para  $\theta(0)$  juntamente com as equações de observação e de estado:

$$Y(t) = h(t) [\theta(t), v(t)],$$

$$\theta(t) = g(t) [\theta(t-1), w(t)].$$

para funções arbitrárias  $g(t)$  e  $h(t)$ . Os modelos de espaço de estado linear especificam essas funções como lineares, e os modelos lineares gaussianos acrescentam as suposições de normalidade para os distúrbios  $v(t)$  e  $w(t)$ . A hipótese de normalidade é plausível em muitas aplicações.

### 3.2 Modelo fatorial dinâmico

Para estimação do modelo, foram aplicados aqui os procedimentos de inferência bayesiana – Kim e Nelson (1999), West e Harrison (1997), Carter e Kohn (1994), Gamerman e Lopes (2006), Petris, Petrone e Campagnoli (2009), entre outros – que, entre outras vantagens, permitem resultados mais confiáveis em contextos nos quais a dimensão temporal da base de dados é pequena em relação ao espaço da dimensionalidade dos parâmetros – além de outras dificuldades, dificilmente tratáveis por meio de procedimentos *frequentistas* (Gelman *et al.*, 2003, p. 696).<sup>3</sup>

Sejam  $m$  variáveis relacionadas que obedecem à distribuição normal multivariada denotada por  $N(0, \Omega)$ , em que  $\Omega$  denota uma matriz  $m \times m$  não singular de variância-covariância. Uma amostra aleatória de tamanho  $T$  é denotada por  $\{y_t, t = 1, \dots, T\}$ . O modelo utilizado neste estudo é dado por

$$y_t = \beta f_t + S_t + e'_t,$$

$$f_t = X_{1t} \theta_{1t} + X_{2t} \theta_{2t} \dots + X_{kt} \theta_{kt} + v'_t \quad (3)$$

$$\theta_{it} = \Gamma_{i1t} \theta_{it-1} + \Gamma_{i2t} \theta_{it-2} \dots + \Gamma_{iNt} \theta_{it-p} + w_{it}, \quad (4)$$

em que  $y_t$ ,  $f_t$ ,  $e_t$  e  $w_t$  são vetores de dimensão  $m$ ,  $k$ ,  $m$  e  $k$ , respectivamente,  $\beta$  é uma matriz  $m \times k$ ,  $\Gamma_i = \text{diag}(\gamma_{i1}, \gamma_{i2}, \dots, \gamma_{ik})$ ,  $i = 1, \dots, p$ ,  $e_t \sim N(0, \Sigma)$ ,  $v_t \sim N(0, \Delta)$  e  $w_{it} \sim N(0, W_i)$  definido para cada imposto  $i$ .  $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_m^2)$  e  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k)$ .

Para algum inteiro positivo específico  $k \leq m$ , o modelo das equações (3) e (4) relaciona cada  $y_t$  com um vetor  $k$  – dimensional de variáveis aleatórias  $f_t$ , os fatores comuns, que obedecem a uma estrutura auto-regressiva  $AR(p)$ , e um componente sazonal,  $S_t = (S_{1t} \ S_{2t} \ \dots \ S_{mt})^T$  (West e Harrison, 1997), definida por  $S_{it} = \sum_{j=1}^6 \left[ a_{ij} \text{sen}\left(\frac{2\pi jt}{12}\right) + b_{ij} \cos\left(\frac{2\pi jt}{12}\right) \right] = \varphi_i X_{it}$ , em que

$$X_t = \left( \text{sen}\left(\frac{2\pi t}{12}\right) \ \text{sen}\left(\frac{2\pi 2t}{12}\right) \ \dots \ \text{sen}\left(\frac{2\pi 6t}{12}\right) \ \cos\left(\frac{2\pi t}{12}\right) \ \cos\left(\frac{2\pi 2t}{12}\right) \ \dots \ \cos\left(\frac{2\pi 6t}{12}\right) \right)^T \text{ e}$$

$$\varphi_i = (a_{i1} \ a_{i2} \ \dots \ a_{i6} \ b_{i1} \ b_{i2} \ \dots \ b_{i6}), \text{ para } i=1, \dots, m.$$

A ideia que motiva a aplicação da forma de Fourier para tratar sazonalidade está ligada ao fato de que qualquer padrão cíclico pode ser representado em termos da combinação linear de funções periódicas.<sup>4</sup> A motivação para o uso desta abordagem, além da flexibilidade, é a facilidade de se obter interpretação, já que as trajetórias em forma de ondas podem ser associadas aos variados

3. Para mais detalhes, ver Mendonça e Medrano (2015).

4. Uma função  $g(t)$  é dita periódica se para algum inteiro  $p > 1$ , e todo  $n, t > 0$   $g(t+np) = g(t)$ .

## TEXTO para DISCUSSÃO

padrões observados na prática. Portanto, distintamente do que se verifica em grande parte dos trabalhos empíricos, a sazonalidade é modelada endogenamente, diminuindo assim mais uma fonte de erro no modelo.

Em problemas práticos, especialmente para valores grandes de  $m$ , o número de fatores  $k$  será pequeno quando comparado com  $m$ , de tal modo que muito da estrutura de variância-covariância é explicada pelos fatores comuns. As variâncias idiossincráticas,  $\sigma_i^2$  medem a variabilidade residual em cada uma das variáveis dos dados, uma vez que a contribuição dos fatores já foi considerada.

### 3.3 Restrições no modelo fatorial

Um modelo com  $k$  fatores deve apresentar alguma restrição para ser definido como um modelo único, não apresentando problemas de identificação. Há várias maneiras de identificar o modelo adicionando restrições à matriz  $\beta$ . Optou-se por seguir, aqui, a alternativa usada em Geweke e Zhou (1996) e Aguilar e West (2000), entre outros, ao restringir  $\beta$  de modo que seja uma matriz de bloco triangular inferior, assumindo ser de posto completo, com elementos da diagonal principal iguais a um, como apresentado adiante:

$$\beta = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \beta_{2,1} & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & 0 \\ \beta_{k,1} & \beta_{k,2} & \beta_{k,3} & \cdots & 1 \\ \beta_{k+1,1} & \beta_{k+1,2} & \beta_{k+1,3} & \cdots & \beta_{k+1,k} \\ \vdots & \vdots & \vdots & \cdots & \\ \beta_{m,1} & \beta_{m,2} & \beta_{m,3} & \cdots & \beta_{m,k} \end{pmatrix}.$$

A restrição exposta anteriormente garante que a matriz  $\beta$  tenha posto completo  $k$ , resolvendo os problemas de identificação que surgem pela invariância do modelo sob mudanças de locação da matriz  $\beta$ . Ao mesmo tempo, garante que o número de parâmetros livres no tempo  $t$  na representação fatorial não exceda os  $m(m+1)/2$  parâmetros de  $\Omega$  não restrito. Além de resolver os problemas de identificação, essa estrutura de  $\beta$  influencia a interpretação dos fatores e pode impactar no ajuste do modelo e na escolha do  $k$  em particular, mas não tem impacto nas previsões, pois as variâncias e covariâncias entre as séries são independentes da decisão de modelagem.

### 3.4 Especificação das *prioris*

Para completar a especificação do nosso modelo, deve-se atribuir classes de *prioris* para os parâmetros  $f_0$ ,  $\beta$ ,  $\Gamma_{i_i}$ ,  $i=1, \dots, p$ ,  $\Sigma$ ,  $\Lambda$  e  $\varphi_{i_i}$ ,  $i=1, \dots, m$ :

- vetor de fatores comuns inicial

$$f_0 \sim N(m_0, C_0);$$

- elementos não nulos da matriz de cargas  $\beta$

$$\beta_{ij} \sim N(\mu_0, V_0) \text{ para } i > j;$$

- elementos da matriz diagonal de variância-covariância  $\Sigma$

$$\sigma_i^2 \sim IG\left(\frac{n_\sigma}{2}, \frac{n_\sigma S_\sigma}{2}\right), \text{ para } i=1, \dots, m;$$

- elementos das matrizes diagonais dos coeficientes auto regressivos  $\Gamma_1, \Gamma_2, \dots, \Gamma_p$

$$\gamma_{ij} \sim N(m_{i_j}, S_{i_j}), \text{ para } i=1, \dots, p, j=1, \dots, k;$$

- elementos da matriz diagonal de variância-covariância  $\Lambda$

$$\lambda_j \sim IG\left(\frac{n_\lambda}{2}, \frac{n_\lambda S_\lambda}{2}\right), j=1, \dots, k; \text{ e}$$

- elementos dos vetores de coeficientes sazonais  $\varphi_1, \varphi_2, \dots, \varphi_m$

$$a_{ij} \sim N(m_{s_j}, C_{s_j}) \text{ e } b_{ij} \sim N(m_{s_j}, C_{s_j}), \text{ para } i=1, \dots, m \text{ e } j=1, \dots, 6.$$

### 3.5 Método Markov Chain Monte Carlo (MCMC) para o modelo fatorial com $k$ fatores

O objetivo é obter as *posteriors* dos parâmetros do modelo, assim como fazer inferência sobre o processo fatorial  $\{f_t : t=1, 2, \dots, T\}$ . Como uma solução exata requer o cálculo de integrais intratáveis do ponto de vista analítico, vamos usar o método do amostrador de Gibbs para obter uma aproximação das *posteriors* de interesse. As *posteriors* completas necessárias para implementar o algoritmo MCMC se encontram no apêndice C. Para mais informações sobre como calcular essas condicionais completas no contexto de modelos fatoriais indicamos Kim e Nelson (1999).

### 3.6 Previsão do MFD com componente sazonal

Nesta seção, será descrita a metodologia para previsão  $h$  passos à frente. Para um modelo de ordem de defasagem  $p = 1$ , a extensão para  $p > 1$  ocorre de forma simples e se encontra descrita no apêndice C.

O interesse aqui é na densidade preditiva  $h$  passos à frente:  $p(y_{t+h} | \mathbf{y}) = \int p(y_{t+h} | \underline{f}_{t+h}, \boldsymbol{\beta}, \boldsymbol{\varphi}, \Theta) p(\underline{f}_{t+h} | \underline{f}_t, \boldsymbol{\beta}, \boldsymbol{\varphi}, \Theta) p(\underline{f}_t | \mathbf{y}) d\underline{f}_{t+h} d\underline{f}_t d\boldsymbol{\beta} d\boldsymbol{\varphi} d\Theta$ , em que:  $\Theta = (\boldsymbol{\sigma}, \boldsymbol{\lambda}, \boldsymbol{\gamma})$ ,  $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_m)$ ,  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_{1k})$ ,  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_k)$ ,  $(y_{t+h} | \underline{f}_{t+h}, \boldsymbol{\beta}, \boldsymbol{\varphi}, \Theta) \sim N(\boldsymbol{\beta} \underline{f}_{t+h} + \boldsymbol{\varphi} X_{t+h}, \boldsymbol{\Sigma})$  e  $(\underline{f}_{t+h} | \underline{f}_t, \boldsymbol{\beta}, \boldsymbol{\sigma}) \sim N(\boldsymbol{\mu}_h, \mathbf{V}_h)$ , com  $\boldsymbol{\mu}_h = \boldsymbol{\Gamma}_1^h \underline{f}_t$  e  $\mathbf{V}_h = \sum_{k=1}^h \boldsymbol{\Gamma}_1^{k-1} \boldsymbol{\Lambda} (\boldsymbol{\Gamma}_1^{k-1})^T$ , para  $h > 0$ .

Portanto, se  $\{(\boldsymbol{\beta}^{(1)}, \underline{f}_t^{(1)}, \boldsymbol{\varphi}^{(1)}, \Theta^{(1)}), \dots, (\boldsymbol{\beta}^{(M)}, \underline{f}_t^{(M)}, \boldsymbol{\varphi}^{(M)}, \Theta^{(M)})\}$  é uma amostra de  $p(\underline{f}_t, \boldsymbol{\beta}, \boldsymbol{\varphi}, \Theta | \mathbf{y})$ , basta gerar uma amostra de tamanho  $M$  de  $(\underline{f}_{t+h} | \underline{f}_t, \boldsymbol{\beta}, \boldsymbol{\varphi}, \Theta)$ , isto é, gerar  $\{(\underline{f}_{t+h}^{(1)}), \dots, (\underline{f}_{t+h}^{(M)})\}$ , de modo que  $\hat{p}(y_{t+h} | \mathbf{y}) = M^{-1} \sum_{j=1}^M p(y_{t+h} | \underline{f}_{t+h}^{(j)}, \boldsymbol{\beta}^{(j)}, \boldsymbol{\varphi}^{(j)}, \Theta^{(j)})$  seja uma aproximação de Monte Carlo para  $p(y_{t+h} | \mathbf{y})$ . Analogamente, uma amostra  $\{y_{t+h}^{(1)}, \dots, y_{t+h}^{(M)}\}$  de  $p(y_{t+h} | \mathbf{y})$  é obtida apresentando  $y_{t+h}^{(j)}$  de  $p(y_{t+h} | \underline{f}_{t+h}^{(j)}, \boldsymbol{\beta}^{(j)}, \boldsymbol{\varphi}^{(j)}, \Theta^{(j)})$ , para  $j = 1, \dots, M$ .

### 3.7 Comentários sobre a inferência bayesiana

Na estimação econométrica, muitas vezes, surge a questão relacionada ao problema de ordem de integração devido às séries serem não estacionárias. Quanto a esse ponto, deve-se ressaltar que a abordagem bayesiana minimiza o problema da ordem de integração das séries, bem como diminui a relevância do tamanho da amostra. Diversos autores (Dejong e Whiteman, 1991; Koop, 1992; Sims e Uhlig, 1991) advogam vigorosamente em favor da alternativa bayesiana sobre a abordagem clássica mais tradicional usada no que se refere à questão da raiz unitária. Por exemplo, sabe-se que os testes do tipo Dickey-Fuller aumentado (ADF) são de baixa potência frente a alternativas plausíveis, especialmente em relação à hipótese alternativa de tendência estacionária. A abordagem bayesiana, por sua vez, revelaria que as hipóteses de raiz unitária e de tendência estacionária apresentariam probabilidades bastante similares quanto às suas funções *a posteriori*. Assim, a abordagem bayesiana fornece um tratamento mais fidedigno da informação amostral que a abordagem clássica.

Outro problema com os testes clássicos de raiz unitária é a descontinuidade gerada na teoria assintótica (Sims, 1988). A abordagem bayesiana, por ser baseada na função de distribuição *a posteriori*, não apresenta o problema da descontinuidade. Koop (1992) assinala que na abordagem clássica os valores críticos gerados a partir de pequenas amostras podem diferir substancialmente

dos valores críticos assintóticos. A abordagem bayesiana, visto que é condicional à amostra observada, fornece resultados para pequenas amostras mais exatos. Em resumo, o problema da raiz unitária não é um ponto crítico na estatística bayesiana. Por fim, diferentemente da abordagem clássica, a prática bayesiana não é dependente do teorema central do limite, o que elimina o uso de propriedades assintóticas (Gelman *et al.*, 2003, p. 696). Isto traz uma importante vantagem em termos da habilidade de obter estimativas mais confiáveis em um contexto de pequenas amostras e alta dimensionalidade paramétrica.

#### 4 TRATAMENTO DA BASE DE DADOS E ESTRATÉGIA DE MODELAGEM

O Confaz disponibiliza informações mensais sobre o comportamento do comércio entre as UFs. A disponibilização do boletim comercial é com base nos dados das notas fiscais eletrônicas (NFe), que se encontram depositados no *ambiente nacional*. Entre tais informações estão as diversas séries do ICMS, o IPVA, o ITCD etc. O ICMS aparece para diferentes níveis de desagregação, por exemplo, o ICMS arrecadado nos setores primário, secundário, terciário etc. Para cada uma das 27 UFs, temos dezessete séries genuínas perfazendo um total de 442 séries. As séries disponibilizadas pelo Confaz estão listadas no apêndice A.

Entender a relação existente entre arrecadação tributária e ambiente macroeconômico é fundamental para a previsão de receitas, uma vez que a evolução das bases tributárias ou os custos/rendimentos de novas medidas de receitas são usualmente preditas usando indicadores macroeconômicos cujos comportamentos são bem conhecidos. Assim, na elaboração das projeções da arrecadação tributária estadual, utilizam-se os cenários mensais de estimativa do produto interno bruto (PIB) e de Índice Nacional de Preços ao Consumidor Amplo (IPCA).

Mesmo em um modelo multivariado que permite contemplar um elevado número de séries, como é o caso do modelo fatorial, entende-se que não cabe tratar todas as séries em um único modelo. Portanto, é necessário optar por algum tipo de desagregação. Sendo assim, a proposta se dá pela escolha de dois modos de desagregação, com base nos critérios de avaliação dentro da amostra e demais procedimentos estatísticos, para avaliar qual das duas estratégias produz a melhor acurácia para a previsão dentro da amostra. As duas formas de desagregação são as seguintes: i) tem-se que, para cada estado, estima-se um modelo com todos os tributos relativos a esse estado; ou ii) estima-se um modelo com a mesma categoria de tributo de todos os estados. Conforme dito na introdução, não é claro qual das duas formas pode gerar projeções. No entanto, é provável que a agregação pelo mesmo tipo de tributo pode reduzir a heterogeneidade entre as séries.

## 4.1 Estratégia de modelagem

O planejamento tributário contribui para o governo equilibrar o seu orçamento. Para o ente público estatal é de extrema importância estimar a arrecadação de receitas, dando-se ênfase ao ICMS, o qual representa a maior fonte de recursos se comparado com os demais tributos, reduzindo-se assim as incertezas futuras nas contas públicas estaduais.

Apesar da disponibilidade de dados, a depender da UF e imposto, desde janeiro de 2006, por questões práticas, a avaliação da capacidade preditiva da arrecadação para os entes subnacionais por meio do modelo fatorial descrito foi realizada mensalmente entre o período de maio de 2020 a maio de 2021, à medida que se teve acesso aos novos valores de arrecadação. Conforme descrito, a estimação do modelo foi implementada usando duas organizações distintas para a base de dados, a fim de investigar a capacidade preditiva do modelo fatorial, e que serão detalhadas a seguir.

A primeira abordagem foi adotada nos três primeiros períodos, ocasião em que a variável dependente do modelo ( $y$ ), descrita anteriormente, foi utilizada com os valores correntes de arrecadação. Nessa abordagem, as previsões foram feitas por UF, isto é, foi criado um modelo fatorial para predição da arrecadação de todos os impostos de cada estado. Além disso, foram como variáveis exógenas o IPCA e o PIB mensal em valores correntes. Em tal abordagem, todas as séries foram utilizadas sem nenhuma manipulação.

Na segunda abordagem, as previsões foram feitas por impostos, isto é, foi criado um modelo fatorial para predição da arrecadação de um determinado imposto para todos os entes subnacionais. A principal justificativa para mudança da organização das séries temporais é que a dinâmica das séries temporais de um imposto qualquer para todos os estados tende a ser menos heterogênea do que a dinâmica encontrada quando se agregam todas as séries temporais de um determinado imposto. Isso permite que o modelo fatorial tenha menor dificuldade para identificação dos fatores.

Nesta segunda abordagem, todas as séries foram transformadas para o  $\log$ -retorno, que permite obter séries com média constante e é amplamente utilizado em finanças, principalmente em alocação de ativo (Dai *et al.*, 2021) e na modelagem de séries temporais com abordagem bayesiana (Qiu, Jammalamadaka e Ning, 2018; Sun, Lee e Emura, 2020).

Os *outliers* foram tratados nas caudas com truncamento em 0,1%, fazendo com que o modelo utilize 99,8% da abrangência das observações. Assim garantiu-se que o poder de generalização do modelo não seja comprometido pelos valores extremos existentes em séries temporais, tendo em vista que a presença dos *outliers* podem comprometer os procedimentos de análise de séries

temporais (Tsay, 1988). Esse tratamento para os valores foi utilizado diante da existência de inúmeras séries temporais, o que inviabilizaria a análise específica de identificação e tratamento de *outliers*.

Por fim, para facilitar o processo de estimação e não comprometer o modelo diante do efeito da pandemia de covid-19, foi adicionada uma variável tipo *dummy* com valor 1 entre abril de 2020 e maio de 2021. Esse tipo de variável é adotado na literatura, como pode ser visto no trabalho de Arin e Braunfels (2018), e permite capturar uma eventual mudança em nível das séries relacionadas à arrecadação de impostos dos entes subnacionais.

A seguir, será apresentado um procedimento estatístico para avaliar a performance preditiva dessas duas alternativas com o propósito de escolher qual delas é a melhor.

## 4.2 Validação da projeção

Nesta seção será apresentado o impacto do poder preditivo da mudança metodológica. Em seguida, serão mostrados os valores preditos para a arrecadação dos entes subnacionais.

Para avaliar o impacto da mudança metodológica serão comparadas as medidas de qualidade após a utilização das bases de dados por impostos. Foram usados diferentes critérios concorrentes para avaliar a capacidade preditiva do modelo. São eles:

- a raiz quadrada do erro quadrático médio (RMSE)<sup>5</sup> que penaliza os *outliers*;
- o desvio absoluto médio (*mean absolute deviation* – MAD);<sup>6</sup>
- o erro percentual absoluto médio (*mean absolute percentage error* – MAPE);<sup>7</sup> e
- o erro percentual médio (*mean percentage error* – MPE).<sup>8</sup>

Por meio da tabela 1, percebem-se os indicadores de qualidade do modelo por imposto entre os períodos de maio de 2020 e maio de 2021.

5.  $RMSE = \sqrt{N^{-1} \sum_{i=1}^N e_i^2}$  em que  $e_i = Y_i - \hat{F}_i$ , em que  $Y_i$  é o valor observado e  $\hat{F}_i$  o valor previsto.

6.  $MAD = N^{-1} \sum_{i=1}^N |e_i|$ .

7.  $MAPE = \left[ \frac{1}{n} \sum_{j=1}^n \frac{|e_j|}{|y_j|} \right] \times 100$ .

8.  $MPE = \left[ N^{-1} \sum_{i=1}^N \frac{e_i}{Y_i} \right] \times 100$ .

**TABELA 1**

**Comparação da melhoria das medidas de qualidade por imposto da primeira metodologia<sup>1</sup> e a segunda metodologia<sup>2</sup>**

(Em %)

Imposto	MAD	MAPE	MPE	RMSE
Comércio atacadista	-56,27	-63,24	-85,95	-53,03
Comércio varejista	-51,88	-58,89	-94,56	-45,08
Dívida ativa	47,49	-68,99	-141,18	14,89
Energia elétrica secundária	-70,69	-87,64	-101,63	-91,49
Energia elétrica terciário	-18,1	-71,71	-113,84	-72,31
ICMS primário	-14,29	9,06	-122,83	-21,21
ICMS secundário	-63,49	-66,12	-74,93	-63,06
IPVA	-54,53	-24,6	-156,54	-62,47
ITCD	-42,05	-38,8	-84,28	-36,01
Outras fontes de receitas	8,07	0,94	-90,81	12,15
Outros ICMS	-63,35	-82,42	-107,25	-73,40
Outros tributos	5,75	40,08	-139,48	-5,84
Petróleo, combustível e lubrificantes secundário	-45,55	-42,72	-96,68	-34,71
Petróleo, combustível e lubrificantes terciário	-13,09	-24,10	-246,82	15,69
Serviços de comunicação	-43,11	-42,04	-114,13	-46,49
Serviços de transporte	-69,15	-79,10	-196,74	-70,75
Taxas	-70,55	-53,50	-95,99	-71,88
Total da arrecadação do ICMS	-61,63	-64,07	-81,37	-59,52
Total da energia elétrica	-52,00	-86,89	-103,25	-89,37
Total de petróleo, combustível e lubrificantes	-40,33	-41,69	-87,60	-31,32
Total terciário	-62,90	-66,57	-91,68	-60,66

Elaboração dos autores.

Notas: <sup>1</sup> Modelo com predição por ente subnacional e adotada entre maio de 2020 e julho de 2020.

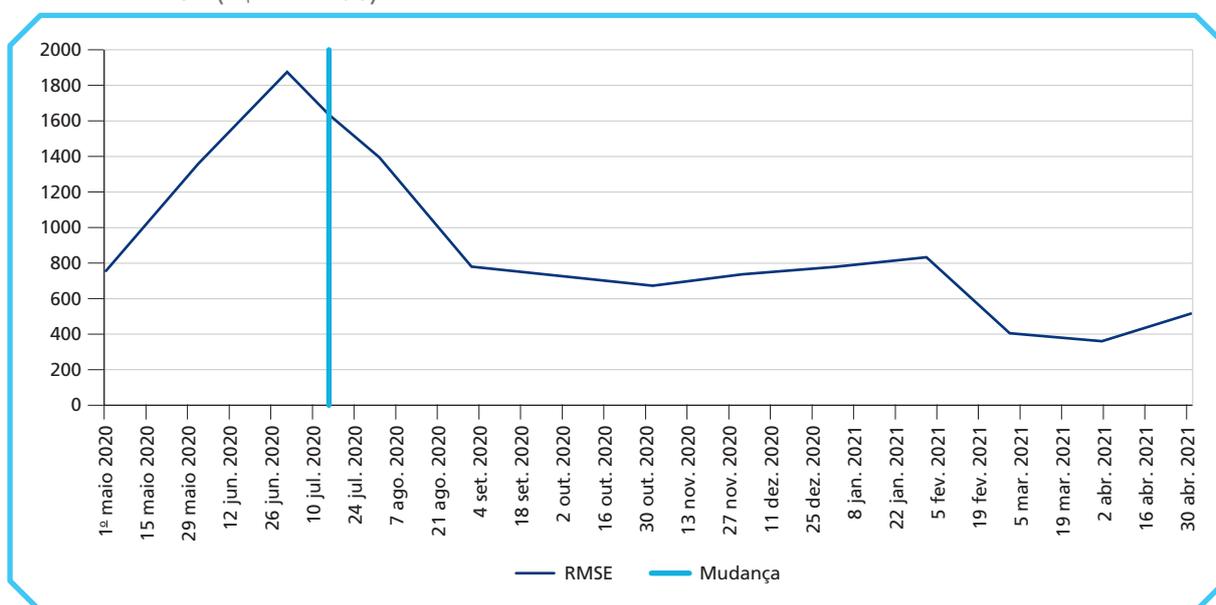
<sup>2</sup> Possui modelo com predição por imposto e adotada entre agosto de 2020 e maio de 2021.

O gráfico 1 mostra o comportamento agregado dos impostos considerando os indicadores de qualidade, antes e depois da mudança de metodologia, representado pela linha vermelha. Nota-se que todos os indicadores apresentam uma melhoria para previsão para os períodos de maio de 2020 a maio de 2021. Conforme dito, nessa nova abordagem, foram introduzidas quatro mudanças metodológicas conforme já descritas. Embora não se possa afirmar categoricamente

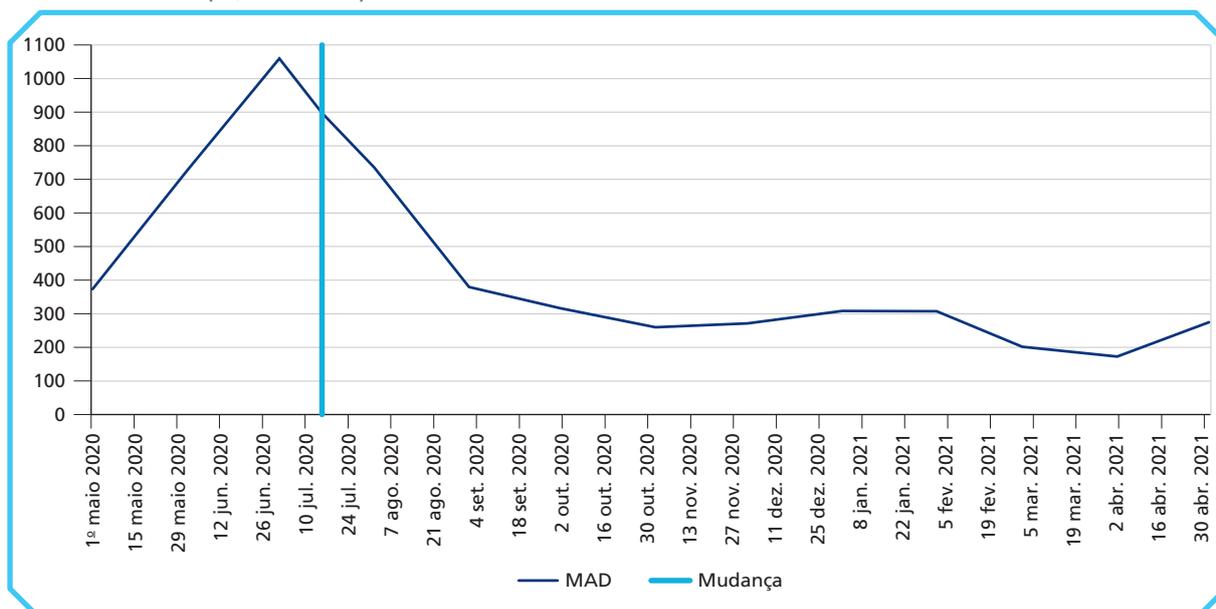
que a segunda estratégia é superior à primeira, cabe dizer que uma análise comparativa feita com base na primeira proposta metodológica (agrupamento por estado) sem e com tratamento dos dados (transformação *log*-retorno, truncamento nas caudas para controle de *outliers* e *dummy* para controle do efeito da covid-19) não apresentou impacto significativo.

### GRÁFICO 1

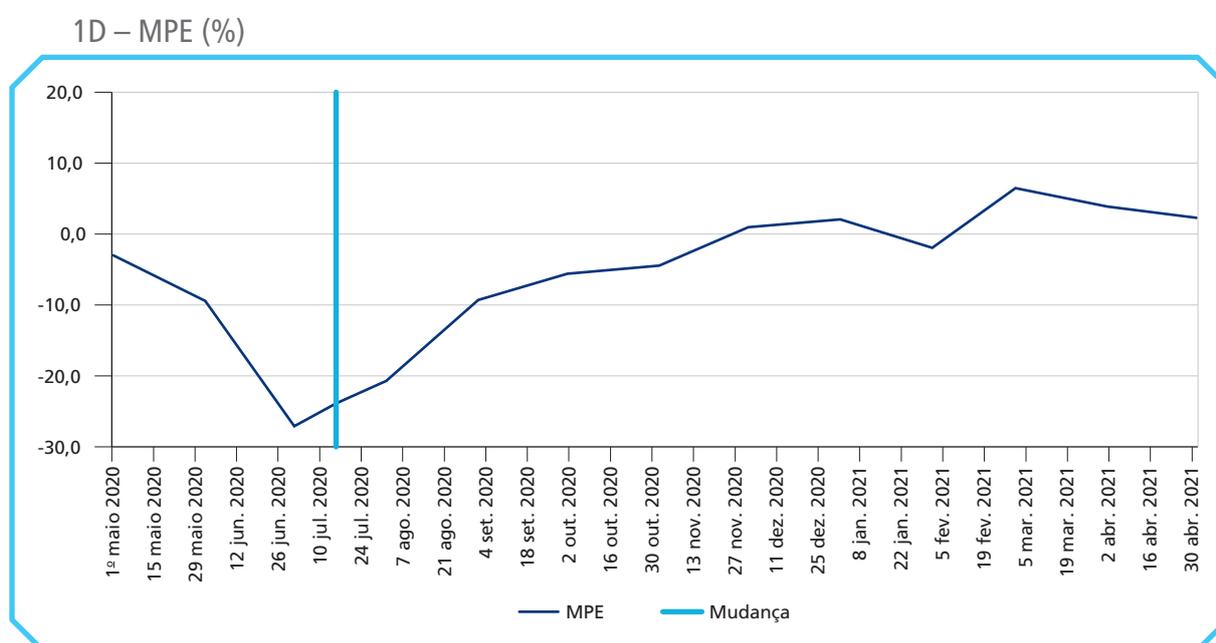
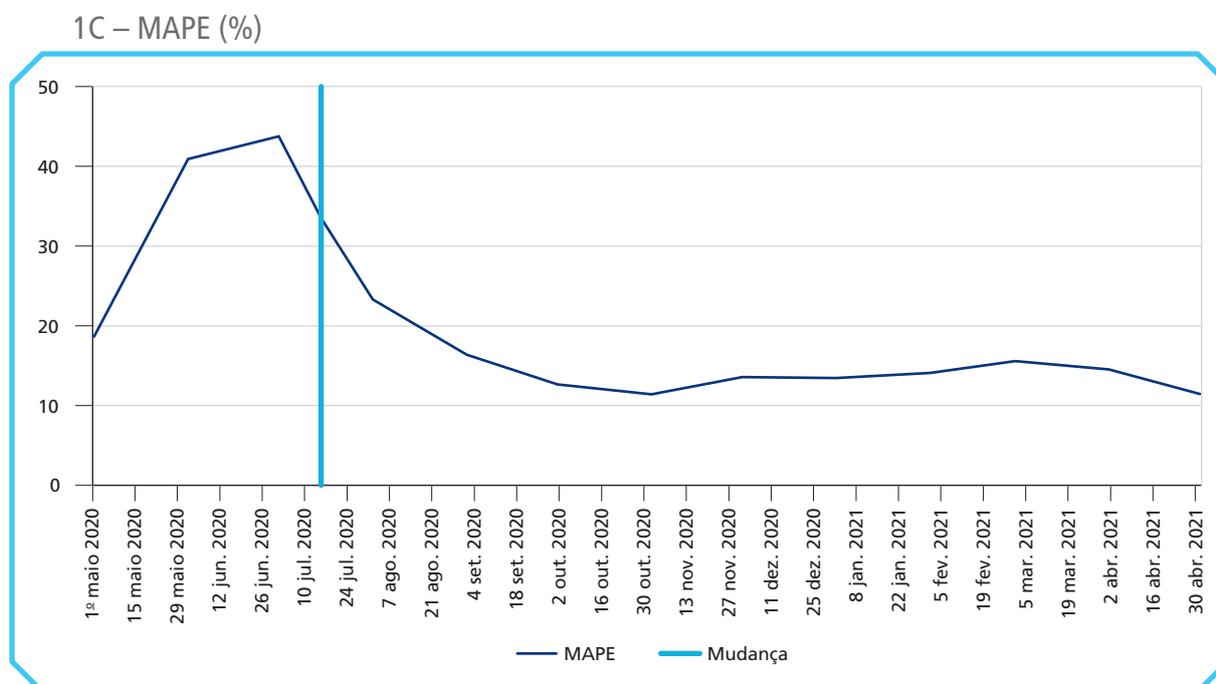
Comportamento agregado dos impostos e medidas de qualidade (1º maio 2020-1º maio 2021)  
1A – RMSE (R\$ 1 milhão)



1B – MAD (R\$ 1 milhão)



## TEXTO para DISCUSSÃO



Elaboração dos autores.

A tabela 2 apresenta o ganho em poder preditivo do modelo por UF. Para o estado de São Paulo percebe-se redução de erro superior a 50% em todos os indicadores.

**TABELA 2**

**Comparação da melhoria das medidas de qualidade por UF da primeira metodologia<sup>1</sup> e a segunda metodologia<sup>2</sup>**

(Em %)

UF	MAD	MAPE	MPE	RMSE
Acre	-51,56	-45,79	-115,5	-48,72
Alagoas	-75,50	-69,72	-93,77	-74,75
Amapá	-45,08	-35,48	-92,58	-36,57
Amazonas	-67,46	-51,62	-99,61	-68,87
Bahia	-63,19	-55,74	-94,67	-63,15
Ceará	-75,49	-66,84	-105,77	-75,30
Distrito Federal	-33,39	-71,81	-97,48	-64,62
Espírito Santo	-58,77	-44,51	-96,82	-58,94
Goiás	-79,41	-80,98	-80,13	-78,81
Maranhão	-43,07	-38,10	-87,71	-40,04
Mato Grosso	-67,98	-88,88	-105,79	-85,15
Mato Grosso do Sul	-67,24	-59,15	-80,07	-66,54
Minas Gerais	-62,18	-57,19	-74,71	-60,87
Pará	-62,29	-70,87	-184,46	-64,51
Paraíba	-65,61	-60,65	-97,62	-63,94
Paraná	-7,38	+1,06	-99,18	-2,72
Pernambuco	-65,18	-47,61	-129,65	-59,52
Piauí	-66,72	-58,61	-105,16	-65,54
Rio de Janeiro	-39,88	-35,92	-66,90	-36,41
Rio Grande do Norte	-39,85	-26,28	-97,75	-38,08
Rio Grande do Sul	-70,69	-63,13	-66,90	-70,67
Rondônia	-45,85	-31,13	-94,38	-40,38
Roraima	-28,58	-85,56	-93,86	-57,22
Santa Catarina	-48,9	-40,57	-64,35	-50,84
São Paulo	-58,81	-54,65	-85,46	-58,40
Sergipe	-79,24	-72,96	-88,71	-78,74
Tocantins	-68,65	-53,52	-100,21	-65,56

Elaboração dos autores.

Notas: <sup>1</sup> Modelo com predição por ente subnacional e adotada entre maio de 2020 e julho de 2020.

<sup>2</sup> Modelo com predição por imposto e adotada entre agosto de 2020 e maio de 2021.

A tabela 3 mostra o erro de previsão da arrecadação um passo à frente acumulado e ponderado da arrecadação para cada tributo, calculado para o período de abril de 2020 a 2021.

## TEXTO para DISCUSSÃO

A ponderação é feita com base na representatividade da arrecadação por categoria de tributo frente ao total de cada estado. Para efeito de comparação, o erro de previsão é também calculado para os dois tipos de desagregação discutidos anteriormente.

A tabela 4 apresenta o mesmo indicador, mas segregando por estado. Conforme se pode observar, há um ganho substancial de desempenho na previsão advindo em estimar o modelo fatorial por categoria de tributo, pois para a maioria dos tributos ocorre uma forte redução do erro de previsão. Para o total da arrecadação, o erro de previsão vai de 8,68% para 0,56%.

**TABELA 3**  
**Erro de previsão da arrecadação tipo de tributo**  
(Em %)

Imposto	Antes	Depois
Comércio atacadista	0,79	-1,53
Comércio varejista	10,13	-1,43
Dívida ativa	26,83	-15,66
Energia elétrica secundária	15,01	-0,60
Energia elétrica terciário	6,73	-2,45
ICMS primário	-0,33	4,74
ICMS secundário	5,45	-2,72
IPVA	13,83	10,76
ITCD	4,12	0,42
Outras fontes de receitas	9,86	-3,55
Outros ICMS	4,26	-1,01
Outros tributos	7,66	2,24
Petróleo, combustível e lubrificantes secundário	8,89	-0,51
Petróleo, combustível e lubrificantes terciário	19,65	-1,96
Serviços de comunicação	5,03	-0,79
Serviços de transporte	92,20	-1,78
Taxas	9,77	-0,82
Total da arrecadação do ICMS	6,37	-1,74
Total da energia elétrica	6,54	-1,47
Total de petróleo, combustível e lubrificantes	10,57	-0,77
Total terciário	4,86	-1,50
Total da arrecadação	8,68	-0,58

Elaboração dos autores.

**TABELA 4**  
**Erro de previsão por estado**  
 (Em %)

UF	Antes	Depois
Acre	13,64	1,93
Alagoas	2,16	0,24
Amapá	17,76	-6,05
Amazonas	3,37	-1,81
Bahia	8,67	-0,45
Ceará	10,10	0,47
Distrito Federal	8,28	-9,02
Espírito Santo	-0,75	0,40
Goiás	-17,37	-0,92
Maranhão	8,46	-1,31
Mato Grosso	14,33	0,02
Mato Grosso do Sul	-1,55	-0,66
Minas Gerais	4,96	-0,63
Pará	-0,52	1,45
Paraíba	2,03	-0,29
Paraná	11,04	1,95
Pernambuco	13,05	4,44
Piauí	6,21	1,64

Elaboração dos autores.

O comportamento da arrecadação da receita tributária e as mudanças nas variáveis macroeconômicas são relevantes para a condução das finanças públicas dos entes subnacionais. Neste contexto, é fundamental para o gestor público ter uma metodologia robusta para projetar a arrecadação tributária estadual. Por exemplo, o ICMS é a principal fonte de arrecadação tributária dos estados e do Distrito Federal, cuja competência cabe a esses entes subnacionais.

O gráfico 2 mostra a relevância de cinco das principais UFs e cinco dos principais impostos. Depreende-se que há maior relevância para o estado de São Paulo, com quase 30% do que é arrecadado no país, e para o ICMS secundário, respondendo com mais de 20% do que é arrecadado no país. O apêndice B apresenta a tabela B.1, que ilustra a importância de cada categoria de tributo no total da arrecadação.

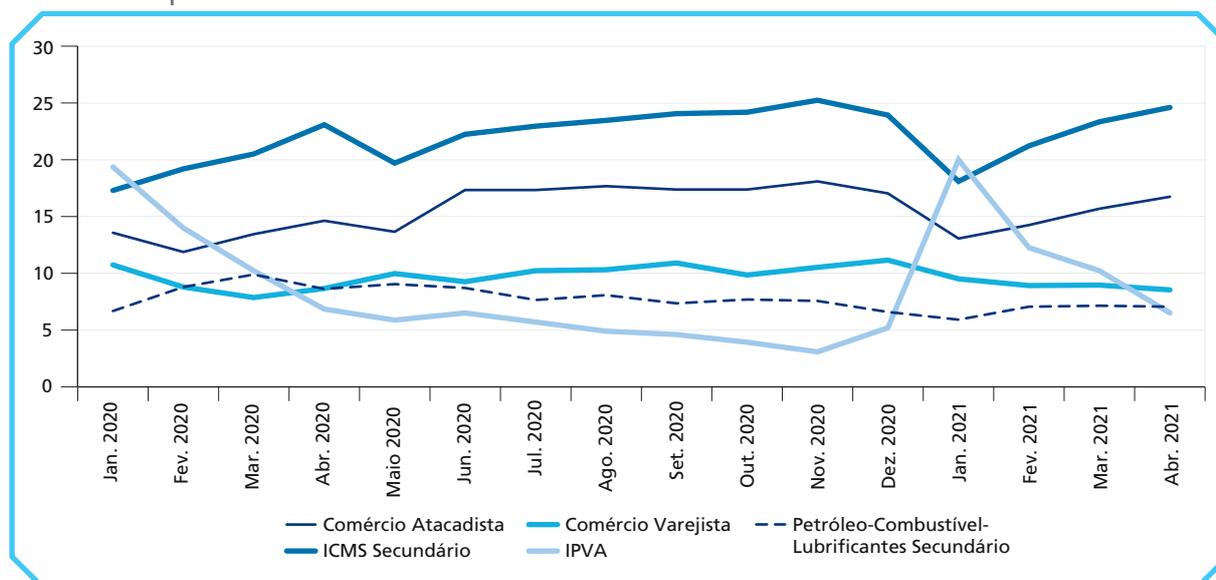
## TEXTO para DISCUSSÃO

### GRÁFICO 2

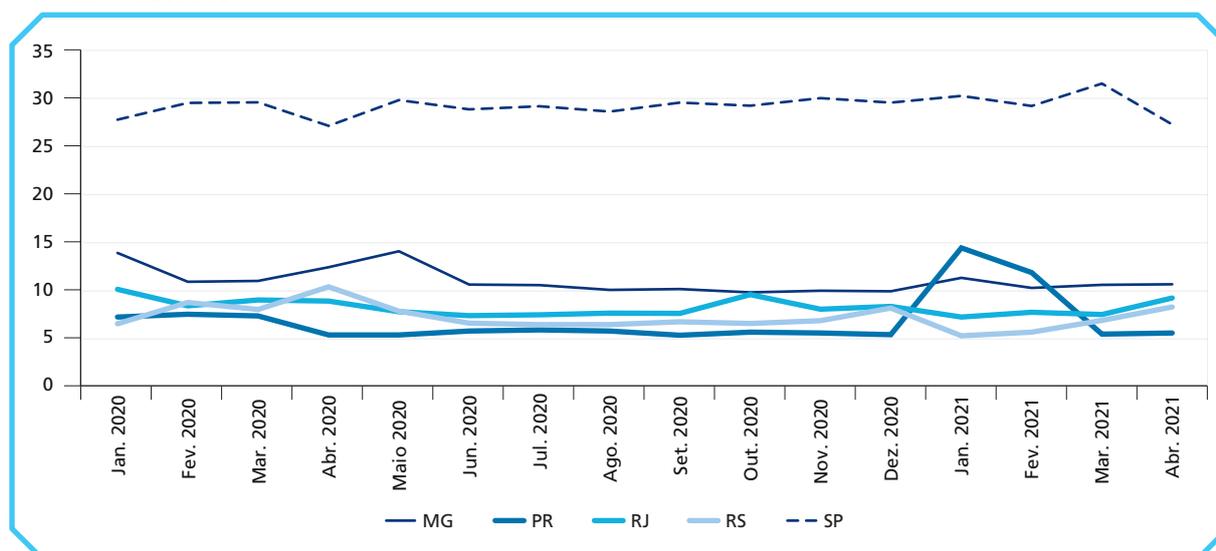
Identificação dos cinco impostos mais relevantes e dos cinco entes federados com maior arrecadação (jan. 2020-abr. 2021)

(Em %)

2A – Impostos



2B – Estados



Elaboração dos autores.

### 4.3 Projeção para arrecadação

Na seção 3 observou-se que o emprego do modelo dinâmico fatorial estimado por categoria de tributo mostra um desempenho preditivo superior àquele em que a amostra é formada tomando

todos os tributos de um mesmo estado. No entanto, a tarefa agora é verificar se para o objetivo em questão, que é o de arrecadação de tributos, a previsão gerada por meio dessa metodologia alcança resultados que de fato possam ser usados pelo gestor de política.

Conforme dito na seção 1.5, o orçamento público por meio da LRF requer que os entes subnacionais apresentem as previsões para as suas receitas de acordo com o que está estabelecido na LDO e na LOA. Assim, de modo a fazer frente a essas necessidades, é necessário que entes federativos disponham de instrumentos capazes de projetarem a arrecadação de modo razoavelmente próximo à realidade. Cabe ainda ressaltar que os órgãos arrecadadores estão mais interessados na previsão de arrecadação acumulada para horizonte temporal mais extenso. Portanto, um erro de previsão negativo em determinado período pode ser compensado por um erro positivo em outro período, de modo que na média seja gerado um resultado satisfatório.

Do que foi dito depreende-se que é necessária uma avaliação do desempenho preditivo para um período mais longo com um indicador adequado. Na tabela 5 é apresentado o resultado da avaliação da previsão quando projetada para um período de seis meses. O erro acumulado percentual está calculado tanto para categoria de tributo como por ente federativo. Tomando-se por corte um erro de previsão de 10%, verifica-se que a metodologia empregada demonstra resultados satisfatórios. Para a maior parte dos tributos, o erro de previsão se situou dentro dessa faixa, sendo que esse erro para o total da arrecadação se situou em -10,20%. Desnecessário dizer que o cálculo do erro é ponderado pela representatividade do tributo na arrecadação total.

**TABELA 5**  
**Avaliação da previsão seis períodos à frente**  
(Em %)  
5A – Tributos

Tributo	Erro percentual
Comércio atacadista	-7,41
Comércio varejista	-10,89
Dívida ativa	-35,97
Energia elétrica secundária	-4,43
Energia elétrica terciário	-10,28
ICMS primário	3,13
ICMS secundário	-15,19
IPVA	2,13
ITCD	-10,51
Outras fontes de receitas	7,93

(Continua)

## TEXTO para DISCUSSÃO

(Continuação)

Tributo	Erro percentual
Outros ICMS	-4,19
Outros tributos	-52,35
Petróleo, combustível e lubrificantes secundário	-6,98
Petróleo, combustível e lubrificantes terciário	0,89
Serviços de comunicação	-0,19
Serviços de transporte	-6,98
Taxas	-11,97
Total da arrecadação do ICMS	-8,26
Total da energia elétrica	-6,81
Total de petróleo, combustível e lubrificantes	-7,07
Total terciário	-5,56
Total da arrecadação	-10,20

### 5B – UFs

UF	Erro percentual
Acre	4,89
Alagoas	-9,65
Amapá	-21,64
Amazonas	-4,44
Bahia	-5,82
Ceará	2,80
Distrito Federal	-28,97
Espírito Santo	1,07
Goiás	-8,74
Maranhão	0,16
Mato Grosso	-7,78
Mato Grosso do Sul	-1,85
Minas Gerais	-10,54
Pará	5,37
Paraíba	-6,85
Pernambuco	9,79
Paraná	-37,16
Piauí	-10,34
Rio de Janeiro	-10,20
Rio Grande do Norte	-4,76

(Continua)

(Continuação)

UF	Erro percentual
Rio Grande do Sul	-15,20
Rondônia	-8,03
Roraima	12,10
Santa Catarina	-2,82
São Paulo	-8,26
Sergipe	-9,42
Tocantins	4,22

Elaboração dos autores.

Uma vez demonstrada a eficácia da metodologia aqui proposta para previsão da arrecadação, deve-se agora discorrer sobre o seu comportamento futuro com base nas projeções. Os avanços na arrecadação do ICMS são reflexo do movimento do nível de atividade econômica, em conformidade com o comportamento de determinados indicadores econômicos. Por exemplo, em 2020, no contexto da crise sanitária ocasionada pela pandemia da covid-19 (coronavírus), o pagamento do auxílio emergencial e o afrouxamento das medidas de isolamento social estimularam o consumo das famílias, refletindo no aumento da arrecadação do ICMS, conforme retratado no desempenho do comércio varejista. A partir de dados mensais realizados de arrecadação de ICMS obtidos no boletim de arrecadação dos tributos estaduais do Confaz,<sup>9</sup> foram elaboradas projeções desse tributo de maio de 2021 até dezembro de 2025. As projeções foram feitas com base no modelo dinâmico fatorial com base no agrupamento por categoria de tributos.

Conforme demonstrado na tabela 6, observa-se que, a partir de 2021, a maior parte dos entes subnacionais conseguem ter projeções de arrecadação de ICMS em patamares acima dos valores observados em 2019. Todavia, cabe considerar que as projeções para o ano de 2021 provavelmente estejam prejudicadas devido à segunda onda da pandemia da covid-19, que gerou um *lockdown*, pelo menos parcialmente, em boa parte do Brasil. Contudo, deve-se levar em conta fatores importantes. Primeiro, não ocorreu uma paralisação completa das atividades, tal como observado em 2020, e nem houve indícios que isso aconteceria. No cômputo geral, a paralisação foi parcial. Segundo, terminado o período de quarentena extrema ao final de 2020, a atividade econômica apresentou recuperação rápida. Terceiro, a expectativa é que a vacinação da população em massa se amplie, fazendo retroceder a incidência dos casos de covid-19. Esses fatores possibilitaram crer que, a despeito da gravidade da situação, a expectativa quanto à confiabilidade das projeções para arrecadação de ICMS aqui apresentadas é favorável na medida em que a atividade

9. Disponível em: <<https://bit.ly/3OGLgnl>>.

econômica deve se recuperar com rapidez findo o período de paralisação. Como meio de controlar o efeito da mudança estrutural gerada no modelo pela pandemia da covid-19, foi incluída como variável exógena, além do PIB e IPCA, uma *dummy* que está presente de agosto de 2020, mês que se iniciou o *lockdown* no Brasil, até o final de 2021.

A arrecadação tributária de *outros tributos* estaduais também ficou afetada pela crise sanitária de 2020. Contudo, os dados reportados na tabela 7 mostram, de maneira similar à arrecadação do ICMS, um melhor resultado de projeções a partir de 2021 para a maioria dos entes subnacionais, refletindo a retomada do nível de atividade econômica no período pós-pandemia.

### TABELA 6

#### Projeções anuais da arrecadação total de ICMS – estados e Distrito Federal (Em R\$ 1 milhão)

UFs e macrorregiões	Realizado		Projeções			
	2019	2020	2021	2022	2023	2024
Acre	1.413	1.374	1.682	2.005	2.618	3.250
Amapá	945	1.016	1.239	1.388	15.773	1.755
Amazonas	10.039	10.807	12.455	14.443	16.792	20.633
Pará	12.250	13.837	15.917	17.517	19.611	22.119
Rondônia	4.013	4.377	5.225	5.528	6.042	6.621
Roraima	1.117	1.240	1.423	1.646	1.940	2.333
Tocantins	3.020	3.295	3.896	4.320	4.845	5.392
Subtotal – Região Norte	32.797	35.946	41.838	46.846	53.420	62.102
Alagoas	4.208	4.369	5.134	5.567	6.078	6.639
Bahia	24.718	24.979	29.377	31.152	33.707	36.741
Ceará	13.152	13.229	14.863	15.718	17.061	18.498
Maranhão	7.883	8.170	9.806	10.746	11.607	13.133
Paraíba	5.904	6.108	7.041	7.924	9.422	11.774
Pernambuco	17.939	17.673	20.072	24.714	33.443	47.835
Piauí	4.489	4.730	5.276	5.823	6.594	7.542
Rio Grande do Norte	5.725	583	6.591	7.442	8.373	9.750
Sergipe	3.548	3.526	4.149	4.535	4.979	5.520
Subtotal – Região Nordeste	87.564	88.617	102.308	113.620	131.265	157.430
Goiás	17.126	17.922	23.578	26.742	30.491	34.990
Distrito Federal	8.182	8.681	9.406	9.996	10.816	11.740
Mato Grosso	13.366	15.335	20.114	23.009	26.033	29.702

(Continua)

(Continuação)

UFs e macrorregiões	Realizado		Projeções			
	2019	2020	2021	2022	2023	2024
Mato Grosso do Sul	10.048	11.095	13.792	15.761	17.768	19.964
Subtotal – Região Centro-Oeste	48.722	53.032	66.890	75.508	85.107	96.395
Espírito Santo	11.452	12.002	14.836	16.179	17.383	18.675
Minas Gerais	51.945	52.461	64.794	70.560	77.006	84.733
Rio de Janeiro	37.015	39.054	53.557	65.657	80.952	99.010
São Paulo	149.774	149.823	173.435	183.402	197.805	215.492
Subtotal – Região Sudeste	250.187	253.340	306.622	335.798	373.145	417.910
Paraná	31.503	31.518	36.983	42.858	50.552	61.017
Rio Grande do Sul	35.743	36.208	45.701	50.635	55.416	60.279
Santa Catarina	23.276	23.095	29.584	32.933	35.760	39.245
Subtotal – Região Sul	90.522	90.821	112.268	126.426	141.728	160.540
<b>Total</b>	<b>509.792</b>	<b>521.756</b>	<b>629.926</b>	<b>698.198</b>	<b>784.665</b>	<b>894.379</b>

Elaboração dos autores.

**TABELA 7****Projeções anuais da arrecadação total de outros tributos – estados e Distrito Federal**  
(Em R\$ 1 milhão)

UFs e macrorregiões	Realizado		Projeções			
	2019	2020	2021	2022	2023	2024
Acre	88	91	106	130	145	155
Amapá	409	387	465	476	681	931
Amazonas	1.293	1.435	1.550	1.844	2.019	2.472
Pará	3.267	3.400	3.372	3.618	4.091	4.989
Rondônia	1.738	1.743	1.640	1.982	2.238	2.545
Roraima	329	313	362	432	495	606
Tocantins	383	436	421	471	544	665
Subtotal – Região Norte	7.508	7.805	7.917	8.954	10.213	12.362
Alagoas	523	560	725	916	1.052	1.335
Bahia	6.351	24.979	29.377	31.152	33.707	36.741
Ceará	1.624	1.169	1.110	1.140	1.175	1.199
Maranhão	789	790	713	738	841	977
Paraíba	526	544	584	669	805	964

(Continua)

## TEXTO para DISCUSSÃO

(Continuação)

UFs e macrorregiões	Realizado		Projeções			
	2019	2020	2021	2022	2023	2024
Pernambuco	1.769	1.792	1.942	2.402	3.848	6.190
Piauí	661	637	735	710	884	1.109
Rio Grande do Norte	430	421	439	543	642	763
Sergipe	372	352	458	558	657	767
Subtotal – Região Nordeste	13.046	31.244	36.084	38.828	43.612	50.045
Goiás	2.577	2.853	3.545	4.075	5.147	6.646
Distrito Federal	1.673	1.596	1.776	1.590	1.612	1.788
Mato Grosso	3.813	3.773	3.712	5.741	6.778	10.188
Mato Grosso do Sul	1.724	2.141	2.566	3.364	4.111	5.069
Subtotal – Região Centro-Oeste	9.787	10.363	11.599	14.771	17.648	23.692
Espírito Santo	1.627	1.731	1.395	1.461	1.580	1.733
Minas Gerais	9.597	9.653	11.142	11.019	11.132	11.396
Rio de Janeiro	8.967	8.712	11.458	12.006	11.891	12.398
São Paulo	26.895	26.910	29.260	27.888	26.303	25.643
Subtotal – Região Sudeste	47.086	47.007	53.254	52.374	50.905	51.169
Paraná	73.553	7.534	8.484	10.497	12.131	14.765
Rio Grande do Sul	4.154	4.311	5.402	6.238	7.193	8.174
Santa Catarina	4.158	4.008	5.193	5.677	6.341	7.052
Subtotal – Região Sul	15.665	15.853	19.079	22.412	25.665	29.991
<b>Total</b>	<b>93.090</b>	<b>112.271</b>	<b>127.933</b>	<b>137.338</b>	<b>148.043</b>	<b>167.259</b>

Elaboração dos autores.

Resumindo, em relação ao varejo, ao final de 2020 e ao longo de 2021, os sinais foram otimistas e apresentaram crescimento das vendas, favorecidos por diversos fatores, tais como reabertura comercial, flexibilização das medidas de distanciamento social, disseminação do comércio eletrônico (*e-commerce*), substituição do consumo de bens e serviços, reação no setor de serviços, assim como transferências emergenciais de renda. Portanto, o crescimento das vendas do varejo indicou continuidade da recuperação do consumo das famílias, e assim se espera que o varejo continue em linha ascendente diante da recomposição da renda das famílias. Conseqüentemente, tem-se observado uma recuperação da arrecadação do ICMS muito mais robusta e rápida do que o esperado para 2020 em relação à maior parte dos entes da Federação.

## 5 CONSIDERAÇÕES FINAIS E IMPLICAÇÕES DE POLÍTICAS

A previsão de arrecadação tributária é fundamental para o planejamento em qualquer ente da Federação. Sem uma boa noção de sua arrecadação futura, um governo não pode planejar adequadamente suas despesas e administrar suas finanças. Portanto, é essencial que as autoridades governamentais possuam boas metodologias de previsões de receitas. Além disso, há três principais atributos relacionados a uma boa previsão: acurácia, enviesada (no sentido de não ser persistentemente otimista ou pessimista), e racional (no sentido de que as previsões incorporam todas as informações relevantes sobre a inter-relação entre a arrecadação tributária e o ambiente macroeconômico).

No caso dos estados e do Distrito Federal, é importante a ênfase dada na arrecadação do ICMS, pois se trata do imposto com maior relevância se comparado com os demais tributos. A arrecadação do ICMS é relevante em termos de custear políticas públicas, projetos e programas sociais, por exemplo. Assim, previsões de arrecadação do ICMS auxiliam o gestor público em dar maior segurança em sua tomada de decisões quanto à alocação de recursos.

Neste estudo foi proposto a aplicação do MDF para projetar a arrecadação tributária dos entes subnacionais (estados e Distrito Federal), cujos resultados procuram incorporar os três atributos acima mencionados. Além disso, a implementação dessa metodologia se deu conjuntamente com dois procedimentos complementares. Primeiro, o emprego de certas técnicas estatísticas e econométricas para tratamento de problemas adicionais como existência de *outliers* e quebra estrutural devido à pandemia da covid-19. Além disso, foi feita uma ampla investigação no intuito de descobrir qual é a melhor maneira para agrupar os dados.

Conforme mostrado, os resultados obtidos foram bastante razoáveis, ficando o erro percentual acumulado para seis meses em torno de 10%. Isso considerando ainda que a amostra é composta de mais de quatrocentas séries contendo padrões de tendência e sazonalidade bastante díspares, além de problema de quebra estrutural, erro de medida, dados faltantes etc.

Contudo, cabe dizer que se deve dar prosseguimento na tarefa de procurar aprimorar a modelagem de modo a reduzir o erro para um patamar ainda menor. Uma estratégia promissora para isso pode ser engendrada a partir das informações constantes no apêndice B e no gráfico 3, que mostra os valores projetados e observados para três categorias de tributos referentes ao estado de São Paulo. De acordo com o apêndice B, o ICMS referente ao setor secundário é o que mais responde pela arrecadação, com cerca de 30% do total e, conforme mostra o gráfico 3, é exatamente o que obtém a pior projeção, tendo em vista ainda que o estado de São Paulo é o principal responsável pela arrecadação, respondendo por cerca de 30% do total arrecadado. Portanto, ao

## TEXTO para DISCUSSÃO

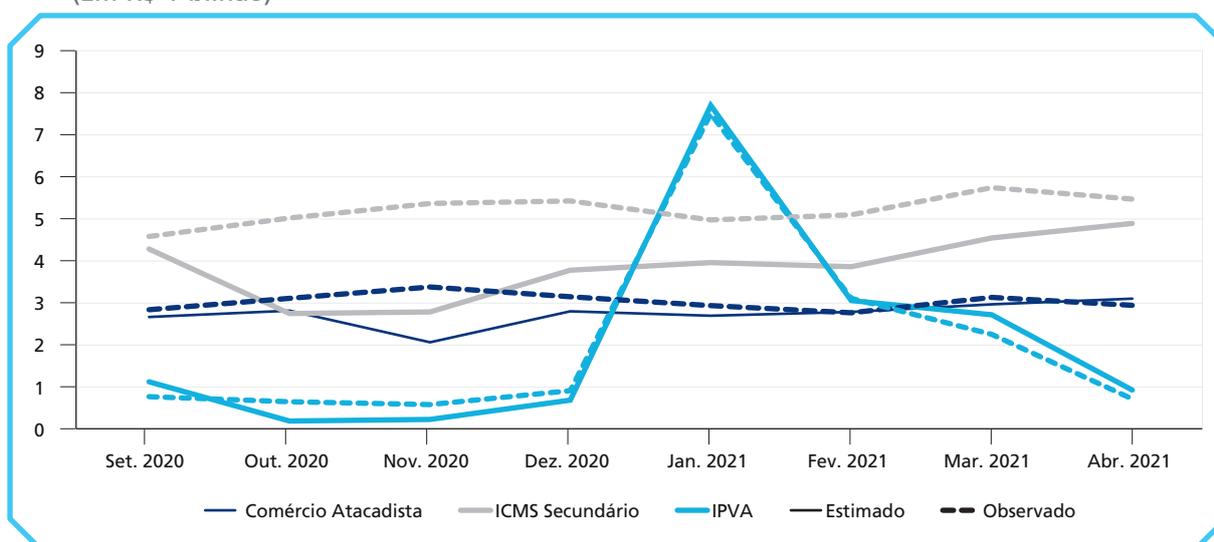
melhorar a previsão desse tributo, tem-se como consequência um substancial aprimoramento na capacidade preditiva do modelo.

Tendo por base que o MDF é estimado para cada categoria de tributo, pode-se então introduzir como variáveis exógenas aquelas que possam representar o fato gerador do tributo com a ressalva de que se deve ter, para tais variáveis exógenas, seus valores observados e preditos, o que não é tão simples.

### GRÁFICO 3

Projetado *versus* observado (set. 2020-abr.2021)

(Em R\$ 1 bilhão)



Elaboração dos autores.

Finalmente, esse estudo apresenta uma proposta metodológica de projeções de arrecadação tributária estadual que beneficiará os gestores públicos no planejamento fiscal de médio prazo.

## REFERÊNCIAS

- AGUILAR, O.; WEST, M. Bayesian dynamic factor models and portfolio allocation. **Journal of Business & Economic Statistics**, v. 18, n. 3, p. 338-357, July 2000.
- ARIN, K. P.; BRAUNFELS, E. The resource curse revisited: a bayesian model averaging approach. **Energy Economics**, v. 70, p. 170-78, 2018.
- AZEVEDO, R. R. de; SILVA, J. M. da; GATSIOS, R. C. 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.
- BARTHOLOMEW, D. J. (Org.). **Latent variable models and factor analysis**. London: Charles Griffin & Co., 1987.
- BATES, J. M.; GRANGER, C. W. J. The combination of forecasts. **Operational Research**, v. 20, n. 4, p. 451-468, Dec. 1969.
- BENELLI, F. C. **Previsão da receita tributária federal por base de incidência**. 2013. Dissertação (Mestrado) – Departamento de Economia, Universidade de Brasília, Brasília, 2013.
- BRASIL. Lei Complementar nº 101, de 4 de maio de 2000. Estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal e dá outras providências. **Diário Oficial da União**, Brasília, 5 mai. 2000.
- \_\_\_\_\_. Ministério da Fazenda. Secretaria do Tesouro Nacional. **Manual de Receita Nacional: aplicado à União, estados, Distrito Federal e Municípios a partir da elaboração e execução da lei orçamentária de 2009**. 1. ed. Brasília: STN, 2008.
- CAMARGOS, A. A. B. **Modelos de previsão da arrecadação tributária do estado de São Paulo: ICMS, IPVA, ITCMD e taxas**. 2008. Dissertação (Mestrado) – Instituto Brasileiro de Mercado de Capitais, São Paulo, 2008.
- CAMPOS, C. V. C. **Previsão da arrecadação de receitas federais: aplicações de séries temporais para o estado de São Paulo**. 2009. Dissertação (Mestrado) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2009.
- CARTER, C. K.; KOHN, R. On Gibbs sampling for state space models. **Biometrika**, v. 81, n. 3, p. 541-553, 1994.

CASTANHO, B. J. da S. **Modelos para previsão de receitas tributárias**: o ICMS do estado do Espírito Santo. 2011. Dissertação (Mestrado) – Centro de Ciências Jurídicas e Econômicas, Universidade Federal do Espírito Santo, Vitória, 2011.

CHAIN, C. P. *et al.* Contribuição da modelagem de valores atípicos na previsão da arrecadação ICMS do estado de Minas Gerais. *In*: SEMINÁRIO EM ADMINISTRAÇÃO, 18., 2015, São Paulo. **Anais...** FEA/USP, 2015.

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.

CORVALÃO, E. D. **Previsão da arrecadação do ICMS em Santa Catarina**: aplicação da abordagem geral para específico em modelos dinâmicos. 2002. Dissertação (Mestrado) – Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2002.

DAI, M. *et al.* A dynamic mean-variance analysis for log returns. **Management Science**, v. 67, n. 2, p. 1093-1108, 2021.

DEJONG, D. N.; WHITEMAN, C. H. The case for trend-stationarity is stronger than we thought. **Journal of Applied Econometrics**, v. 6, n. 4, p. 413-421, 1991.

DUARTE, F. C. de L.; SOUZA, M. F. de; GIRÃO, L. F. de A. P. Previsão da arrecadação do ICMS: uso do modelo Holt-Winters aditivo na Paraíba. *In*: CONGRESSO USP DE INICIAÇÃO CIENTÍFICA EM CONTABILIDADE, 11., 2014, São Paulo. **Anais...** São Paulo, 2014.

DURBIN, J.; KOOPMAN, S. J. (Org.). **Time series analysis by state space methods**. Oxford: Oxford University Press, 2001.

GADELHA, S. R. de B.; LIMA, A. F. R.; POLLI, D. A. Uso da metodologia de combinação de previsões para projeções da arrecadação de receitas brutas primárias de tributos federais. **Cadernos de Finanças Públicas**, Brasília, v. 1, n. 1, Edição Especial, p. 1-70, 2020. Disponível em: <<https://bit.ly/37LXHgY>>. Acesso em: 11 jul. 2021.

GAMERMAN, D.; LOPES, H. F. (Org.). **Markov chain Monte Carlo**: stochastic simulation for bayesian inference. 2nd ed. London: Chapman & Hall/CRC, 2006.

GELMAN, A. *et al.* (Org.). **Bayesian data analysis**. London: Chapman & Hall/CRC, 2003.

GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. **Statistical Science**, v. 7, n. 4, p. 457-472, 1992.

GEWEKE, J.; ZHOU, G. Measuring the pricing error of the arbitrage pricing theory. **The Review of Financial Studies**, v. 9, n. 2, p. 557-587, 1996.

GUARAGNA, P. R. S.; MELLO, M. R. de. **Um modelo de previsão para a arrecadação do ICMS**. 2002. Monografia (Graduação) – Escola de Administração Fazendária, Brasília, 2002.

HAMILTON, J. D. (Org.). **Time series analysis**. New Jersey: Princeton University Press, 1994.

KIM, C.-J.; NELSON, C. R. (Org.). **State-space models with regime switching: classical and Gibbs-Sampling approaches with applications**. Massachusetts: The MIT Press, 1999.

KOOP, G. 'Objective' bayesian unit root tests. **Journal of Applied Econometrics**, v. 7, n. 1, p. 65-82, 1992.

LIMA, E. C. R.; FERREIRA, L. da R.; HONAISSER, V. H. M. B. Projeção das receitas correntes do estado do Rio de Janeiro: aplicações de modelos univariados e multivariados. **(Syn)thesis**, Rio de Janeiro, v. 6, n. 2, p. 207-233, 2013.

MELO, B. S. V. **Modelo de previsão para a arrecadação tributária**. 2001. Monografia (Graduação) – Escola de Administração Fazendária, Brasília, 2001.

MENDONÇA, M. J.; MEDRANO, L. A. **Aplicação do modelo fatorial dinâmico para previsão da receita tributária no Brasil**. Rio de Janeiro: Ipea, 2015. (Texto para Discussão, n. 2064).

MENDONÇA, M. J. C. de; SACHSIDA, A.; MEDRANO, L. A. T. Um modelo econométrico para previsão de impostos no Brasil. **Economia Aplicada**, v. 17, n. 2, p. 295-329, 2013.

MOÇO, J. R. C. **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, Rio de Janeiro, 2017.

OLIVEIRA, W. de. **Curso de responsabilidade fiscal: direito, orçamento e finanças públicas**. 2. ed. Belo Horizonte: Fórum, 2015.

PECEGUINI, E. E. **Análise comparativa de métodos de previsão aplicados à arrecadação do ICMS: estado de São Paulo**. 2001. Monografia (Especialização) – Fundação Instituto de Pesquisas Econômicas, São Paulo, 2001.

PESSOA, F. de M. C.; CORONEL, D. A.; LIMA, J. E. de. 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, p. 47-64, 2013.

PETRIS, G.; PETRONE, S.; CAMPAGNOLI, P. (Org.). **Dynamic linear models with R**. New York: Springer, 2009.

QIU, J.; JAMMALAMADAKA, S. R.; NING, N. Multivariate bayesian structural time series model. **Journal of Machine Learning Research**, v. 19, p. 1-33, 2018.

RUBIN, I. S. **The politics of public budgeting**: getting and spending, borrowing and balancing. 6th ed. Washington: CQPress, 2009.

SIMS, C. A. Bayesian skepticism on unit root econometrics. **Journal of Economic Dynamics and Control**, v. 12, n. 2-3, p. 463-474, 1988.

SIMS, C. A.; UHILG, H. Understanding unit rooters: a helicopter tour. **Econometrica**, v. 59, n. 6, p. 1591-1599, 1991.

SIQUEIRA, M. L. **Melhorando a previsão de arrecadação tributária federal através da utilização de modelos e séries temporais**. 2002. Monografia (Graduação) – Escola de Administração Fazendária, Brasília, 2002.

SUN, L.-H.; LEE, C.-S.; EMURA, T. A bayesian inference for time series via copula-based Markov chain models. **Communications in Statistics – Simulation and Computation**, v. 49, n. 11, p. 2897-2913, 2020.

TSAY, R. S. Outliers, level shifts, and variance changes in time series. **Journal of Forecasting**, v. 7, n. 1, p. 1-20, 1998.

WEST, M.; HARRISON, J. (Org.). **Bayesian forecasting and dynamic models**. New York: Springer, 1997.

## BIBLIOGRAFIA COMPLEMENTAR

FRÜHWIRTH-SCHNATTER, S. Data augmentation and dynamic linear models. **Journal of Time Series Analysis**, v. 15, n. 2, p. 183-202, 1994.

LÜTKEPOHL, H. (Org.). **New introduction to multiple time series analysis**. Berlin: Springer, 2005.

## APÊNDICE A

### BOX A.1

#### Base de dados

Total geral da receita tributária: 1 + 2

1 Total de arrecadação do Imposto sobre a Circulação de Mercadorias e Prestação de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação (ICMS)

1.1 ICMS – Setor Primário (119 – ICMS Primário)

1.2 ICMS – Setor Secundário (129 – ICMS Secundário)

1.3 ICMS – Setor Terciário Total

1.3.1 ICMS – Setor Terciário – Comércio Atacadista (131 – Comércio Atacadista)

1.3.2 ICMS – Setor Terciário – Comércio Varejista (132 – Comércio Varejista)

1.3.3 ICMS – Setor Terciário – Serviço de Transporte (133 – Serviço de Transporte)

1.3.4 ICMS – Setor Terciário – Serviço de Comunicação (134 – Serviço de Comunicação)

1.3.5 ICMS – Setor Terciário – Outros (135 – Outros ICMS)

1.4 ICMS – Energia Elétrica Total

1.4.1 ICMS – Energia Elétrica – Setor Secundário (211 – Energia Elétrica Secundária)

1.4.2 ICMS – Energia Elétrica – Setor Terciário (212 – Elétrica Terciária)

1.5 ICMS – Petróleo, combustíveis e lubrificantes Total

1.5.1 ICMS – Petróleo, combustíveis e lubrificantes – Setor Secundário (221 – Petróleo-Combustível-Lubrificantes)

1.5.2 ICMS – Petróleo, combustíveis e lubrificantes – Setor Terciário (222 – Petróleo-Combustível-Lubrificantes)

1.6 ICMS – Dívida Ativa (239 – Dívida Ativa)

1.7 ICMS – Outras Fontes de Receitas (249 – Outras Fontes de Receitas)

2 Total de arrecadação de outros tributos

2.1 Total de arrecadação do Imposto sobre a Propriedade de Veículos Automotores – IPVA (409 - IPVA)

2.2 Total de arrecadação do Imposto sobre a Transmissão Causa Mortis e Doação de Quaisquer Bens ou Direitos – ITCD (419 – ITCD)

2.3 Total de arrecadação de Taxas (429 – Taxas)

2.4 Total de arrecadação de outros (439 – Outros Tributos)

Elaboração dos autores.

## APÊNDICE B

TABELA B.1

**Representatividade no total da arrecadação por categoria de tributo**

Imposto	Total	%
ICMS Secundário	4.590.211.974	30,18
Comércio atacadista	2.801.817.928	18,42
IPVA	1.476.376.689	9,71
Comércio varejista	1.140.789.329	7,50
Energia elétrica terciário	1.039.508.384	6,83
Petróleo, combustível e lubrificantes secundário	893.731.949	5,88
Serviços de comunicação	799.943.776	5,26
Outras fontes de receitas	747.217.575	4,91
Taxas	510.479.882	3,36
Serviços de transporte	304.469.708	2,00
ITCD	276.726.619	1,82
Petróleo, combustível e lubrificantes terciário	259.514.156	1,71
Outros ICMS	247.415.622	1,63
Dívida ativa	64.273.572	0,42
ICMS primário	43.694.345	0,29
Outros tributos	10.769.340	0,07
Energia elétrica secundária	3.689.003	0,02

Elaboração dos autores.

## APÊNDICE C

## DESCRÇÃO DO ALGORITMO MCMC PARA O MODELO FATORIAL DINÂMICO

Podemos representar o modelo (1) em uma estrutura de espaço de estados, do seguinte modo:

$$\begin{aligned} (y_t - \boldsymbol{\varphi} X_t) &= F^E f_t^E + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim N(0, \boldsymbol{\Sigma}) \\ f_t^E &= G^E f_{t-1}^E + \boldsymbol{w}_t^E, \quad \boldsymbol{w}_t^E \sim N(0, Q^E) \end{aligned} \quad (2)$$

em que:

$$G_{pkxp}^E = \begin{bmatrix} \boldsymbol{\Gamma}_1 & \boldsymbol{\Gamma}_2 & \cdots & \boldsymbol{\Gamma}_{p-1} & \boldsymbol{\Gamma}_p \\ I_{kxk} & \mathbf{0}_{kxk} & \cdots & \mathbf{0}_{kxk} & \mathbf{0}_{kxk} \\ \mathbf{0}_{kxk} & I_{kxk} & \cdots & \mathbf{0}_{kxk} & \mathbf{0}_{kxk} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0}_{kxk} & \mathbf{0}_{kxk} & \cdots & I_{kxk} & \mathbf{0}_{kxk} \end{bmatrix}, \quad \tilde{f}_t^E = \begin{bmatrix} f_t \\ f_{t-1} \\ \vdots \\ f_{t-p+1} \end{bmatrix}, \quad \boldsymbol{w}_t^E = \begin{bmatrix} \boldsymbol{w}_t \\ \mathbf{0}_{kx1} \\ \vdots \\ \mathbf{0}_{kx1} \end{bmatrix},$$

$$Q_{pkxp}^E = \begin{bmatrix} \Lambda_{kxk} & \mathbf{0}_{kxk(p-1)} \\ \mathbf{0}_{k(p-1)xk} & \mathbf{0}_{k(p-1)xk(p-1)} \end{bmatrix}, \quad F_{mxpk}^E = [\boldsymbol{\beta}_{mxk} \quad \mathbf{0}_{mxk} \quad \mathbf{0}_{mxk} \quad \cdots \quad \mathbf{0}_{mxk}].$$

## 1 POSTERIORIS COMPLETAS NECESSÁRIAS PARA A IMPLEMENTAÇÃO DO AMOSTRADOR DE GIBBS

Para  $i = 1, \dots, k$ , escreva  $\boldsymbol{\beta}_i = (\beta_{i1}, \dots, \beta_{ii})^T$  que são os elementos diferentes de zero. Para os restantes das linhas  $i = k + 1, \dots, m$ , escreva  $\boldsymbol{\beta}_i = (\beta_{i1}, \dots, \beta_{ik})$ . Similarmente, para  $i = 1, \dots, k$  denote  $\mathbf{F}_i$  a matriz  $T \times i$  contendo as primeiras  $i$  colunas de  $\mathbf{F}$  e para todo  $i$  seja  $y_i$  a  $i$ -ésima coluna de  $y$ ,  $\mathbf{S}_i$  a  $i$ -ésima coluna de  $(\mathbf{X}\boldsymbol{\varphi}^T)$  com  $\mathbf{X} = (X_1^T, \dots, X_T^T)^T$  e  $\mathbf{y}_i^* = \mathbf{y}_i - \mathbf{S}_i$ . Seja  $\boldsymbol{\gamma}_j = (\gamma_{1j}, \gamma_{2j}, \dots, \gamma_{pj})^T$  o vetor contendo o  $j$ -ésimo elemento da diagonal para cada uma das  $p$  matrizes  $\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \dots, \boldsymbol{\Gamma}_p$ . Defina também para  $j = 1, \dots, k$ ,

$$\boldsymbol{\mathfrak{S}}_j = \begin{pmatrix} f_{p,j} & f_{p-1,j} & \cdots & f_{1,j} \\ f_{p+1,j} & f_{p,j} & \cdots & f_{2,j} \\ \vdots & \vdots & \vdots & \vdots \\ f_{T-1,j} & f_{T-2,j} & \cdots & f_{T-p,j} \end{pmatrix} \text{ e } \mathbf{H}_j = \begin{pmatrix} f_{p+1,j} \\ f_{p+2,j} \\ \vdots \\ f_{T,j} \end{pmatrix}.$$

Para amostrar da condicional completa de  $f_t^E$  siga os passos a seguir.

## TEXTO para DISCUSSÃO

- 1) Aplicar filtro de Kalman para calcular  $m_t = E[f_t^E | D_t]$  e  $C_t = V[f_t^E | D_t]$  para  $t = 1, \dots, T$ , do seguinte modo:

$$a_t = G^E m_{t-1}, R_t = G^E C_{t-1} G^{E^T} + Q^E,$$

$$f_t = F^E a_t, Q_t = F^E R_t F^{E^T} + \Sigma,$$

$$A_t = R_t F^{E^T} Q_t^{-1}, e_t = (y_t - S_t) - f_t$$

e armazenar  $m_t$  e  $C_t$ ,  $t = 1, \dots, T$ . No final use  $m_t$  e  $C_t$  para gerar  $f_t^E | \cdot \sim N(m_t, C_t)$ .

- 2) Defina  $f_t^{E*}$  como sendo os primeiros  $k$  elementos de  $f_t^E$ ,  $G^{E*}$  como sendo as primeiras  $k$  linhas de  $G^E$  e  $Q^{E*}$  como sendo o primeiro bloco  $k \times k$  da matriz  $Q^E$ . Para  $t = T-1, T-2, \dots, 1$  faça:

$$m_t^* = m_t + C_t G^{E*T} (G^{E*} C_t G^{E*T} + Q^{E*})^{-1} (f_{t+1}^{E*} - G^{E*} m_t),$$

$$C_t^* = C_t - C_t G^{E*T} (G^{E*} C_t G^{E*T} + Q^{E*})^{-1} G^{E*} C_t,$$

e gere  $f_t^E | \cdot, f_{t+1}^{E*} \sim N(m_t^*, C_t^*)$  e mantenha somente os  $k$  primeiros elementos gerados dessa forma. Esses primeiros  $k$  elementos formam uma amostra de  $f_t$  da condicional completa.

A *posteriori* condicional completa de  $\Sigma$  se reduz a um conjunto de  $m$  gamas inversas independentes, com  $\sigma_i^2 \sim IG\left(\frac{n_\sigma^*}{2}, \frac{n_\sigma^* S_\sigma^*}{2}\right)$ , em que

$$n_\sigma^* = n_\sigma + T,$$

$$n_\sigma^* S_\sigma^* = n_\sigma S_\sigma + (\mathbf{y}_i^* - \mathbf{F} \beta_i^{*T})^T (\mathbf{y}_i^* - \mathbf{F} \beta_i^{*T}).$$

A *posteriori* condicional completa de  $\beta$  divide-se em:

- para  $i = 1, \dots, k$ ,  $\beta_i \sim N(\underline{\mu}_i, \underline{V}_i) 1(\beta_{ii} = 1)$ , em que  $\underline{\mu}_i = \underline{V}_i (V_0^{-1} \mu_0 1_i + \sigma_i^{-2} \mathbf{F}_i^T \mathbf{y}_i^*)$  e  $\underline{V}_i^{-1} = V_0^{-1} I_i + \sigma_i^{-2} \mathbf{F}_i^T \mathbf{F}_i$ ; e
- para  $i = k+1, \dots, m$ ,  $\beta_i \sim N(\underline{\mu}_i, \underline{V}_i)$ , em que  $\underline{\mu}_i = \underline{V}_i (V_0^{-1} \mu_0 1_k + \sigma_i^{-2} \mathbf{F}_i^T \mathbf{y}_i^*)$  e  $\underline{V}_i^{-1} = V_0^{-1} I_k + \sigma_i^{-2} \mathbf{F}_i^T \mathbf{F}_i$ .

A *posteriori* condicional completa de  $\gamma_j$  é dada por:

- para  $j = 1, \dots, k$ ,  $\gamma_j \sim N(A_j, B_j)$ , em que  $A_j = B_j (S_r^{-1} m_r 1_p + \lambda_j^{-1} \mathfrak{S}_j^T H_j)$  e  $B_j = S_r^{-1} I_{p \times p} + \lambda_j^{-1} \mathfrak{S}_j^T \mathfrak{S}_j$ .

A *posteriori* condicional completa de  $\Lambda$  se reduz a um conjunto de  $k$  gammas

inversas independentes, com  $\lambda_j \sim IG\left(\frac{n_\lambda^*}{2}, \frac{n_\lambda^* S_\lambda^*}{2}\right)$ , em que  $n_\lambda^* = n_\lambda + T$ ,  $n_\lambda^* S_\lambda^* = n_\lambda S_\lambda + (\mathbf{H}_j - \mathfrak{I}_j \underline{\gamma}_j)^T (\mathbf{H}_j - \mathfrak{I}_j \underline{\gamma}_j)$ .

A *posteriori* condicional completa de  $\varphi_i$  é dada por:

para  $i=1, \dots, m$ ,  $\varphi_i \sim N(C_i, D_i)$ , em que  $C_i = D_i(C_s^{-1} m_s \mathbf{1}_{12} + \sigma_i^{-1} \mathbf{X}(y_i - \mathbf{F}\beta_i))$  e  $D_i = C_s^{-1} I_{12 \times 12} + \sigma_i^{-2} \mathbf{X}^T \mathbf{X}$ .

## 2 PREVISÃO MODELO FATORIAL DINÂMICO COM COMPONENTE SAZONAL COM $P > 1$

Passa o modelo (1) para a estrutura de espaço de estados (2) e utilize a mesma metodologia apresentada na seção 2.6, substituindo  $f_t$ ,  $\Lambda$ ,  $\beta$  e  $\Gamma_1$  por  $f_t^E$ ,  $Q^E$ ,  $F^E$  e  $G^E$  respectivamente.

# Ipea – Instituto de Pesquisa Econômica Aplicada

## EDITORIAL

### **Chefe do Editorial**

Aeromilson Trajano de Mesquita

### **Assistentes da Chefia**

Rafael Augusto Ferreira Cardoso

Samuel Elias de Souza

### **Supervisão**

Camilla de Miranda Mariath Gomes

Everson da Silva Moura

### **Revisão**

Alice Souza Lopes

Amanda Ramos Marques

Ana Clara Escórcio Xavier

Clícia Silveira Rodrigues

Idalina Barbara de Castro

Olavo Mesquita de Carvalho

Regina Marta de Aguiar

Reginaldo da Silva Domingos

Brena Rolim Peixoto da Silva (estagiária)

Nayane Santos Rodrigues (estagiária)

### **Editoração**

Anderson Silva Reis

Cristiano Ferreira de Araújo

Danielle de Oliveira Ayres

Danilo Leite de Macedo Tavares

Leonardo Hideki Higa

### **Capa**

Aline Cristine Torres da Silva Martins

### **Projeto Gráfico**

Aline Cristine Torres da Silva Martins

*The manuscripts in languages other than Portuguese published herein have not been proofread.*

## **Missão do Ipea**

Aprimorar as políticas públicas essenciais ao desenvolvimento brasileiro por meio da produção e disseminação de conhecimentos e da assessoria ao Estado nas suas decisões estratégicas.



**ipea** Instituto de Pesquisa  
Econômica Aplicada

