



Seleção de defasagens em testes de raiz unitária: uma revisão de literatura

Lag selection in unit root tests: a literature review

Recebido: 18 abr. 2022

Aprovado: 03 ago. 2022

Versão do autor aceita publicada online: 03 ago. 2022

Publicado online: 12 ago. 2022

Como citar esse artigo - American Psychological Association (APA)

Silveira, A. G., Mattos, V. L. D., Nakamura, L. R., Amaral, M. C., & Konrath, A. C. (jul./set. 2024). Seleção de defasagens em testes de raiz unitária: uma revisão de literatura. *Exacta*, 22(3), p. 852-879.

<https://doi.org/10.5585/exactaep.2022.22061>

Submeta seu artigo para este periódico

Processo de Avaliação: *Double Blind Review*

Editor: Dr. Luiz Fernando Rodrigues Pinto



Dados Crossmark



Seleção de defasagens em testes de raiz unitária: uma revisão de literatura

Lag selection in unit root tests: a literature review

 Anderson Garcia Silveira¹  Viviane Leite Dias de Mattos²  Luiz Ricardo Nakamura³
 Mariane Coelho Amaral⁴ e  Andrea Cristina Konrath⁵

¹ Possui doutorado em andamento pelo Programa de Pós-Graduação em Modelagem Computacional pela Universidade Federal do Rio Grande. Professor do IFRS – Câmpus Rio Grande. Universidade Federal do Rio Grande / Rio Grande, RS -



Brasil

² Doutorado em Engenharia de Produção pela Universidade Federal de Santa Catarina. Professor adjunto da Universidade

Federal do Rio Grande. Universidade Federal do Rio Grande / Rio Grande, RS - Brasil



³ Doutorado em Ciências (Estatística e Experimentação Agronômica) pela Universidade de São Paulo, com período sanduíche na London Metropolitan University. Professor Adjunto do Departamento de Informática e Estatística da Universidade Federal de Santa Catarina. Pesquisador dos grupos GAMLSS (www.gamlss.com) e “Grupo de Análise e

Modelagem Estatística – GAME”. Universidade Federal de Santa Catarina / Florianópolis, SC - Brasil



⁴ Engenheira eletricista. Mestre em Modelagem Computacional. Universidade Federal do Rio Grande / Rio Grande, RS -



Brasil

⁵ Doutorado em Engenharia Mecânica pela Universidade Federal de Santa Catarina. Professora adjunta na Universidade

Federal de Santa Catarina (UFSC). Universidade Federal de Santa Catarina / Florianópolis, SC - Brasil



Nota dos Autores

Autores declaram que não há conflitos de interesses.



Resumo

Dentre as maneiras mais usuais de modelar e realizar previsões de demanda através de séries temporais, como, por exemplo a série de demanda por eletricidade, está a análise econométrica. Neste tipo de análise, a presença de raiz unitária pode fazer com que as previsões se tornem imprecisas. Assim, é fundamental identificar a presença de raiz unitária nas séries a serem modeladas. Para isso, podem ser aplicados os testes de raiz unitária, como por exemplo o teste ADF. Uma das etapas de grande importância na aplicação deste teste é a seleção de defasagens a serem utilizadas. Este trabalho apresenta de maneira qualitativa e quantitativa que a pesquisa por este tema se encontra em fase de expansão e que não existe um consenso no que tange à definição do número de defasagens a serem aplicadas no teste. Assim, fica evidente que este é um assunto em que cabe aprofundamento de pesquisa.

Palavras-chave: raiz unitária, seleção de defasagens, demanda por eletricidade, estacionariedade

Abstract

The econometric analysis may be one of the most common ways to model and forecast different time series problems, such as the electricity demand. In this type of analysis, the presence of unit root may lead to unreliable forecasts. Hence, the correct identification of the presence of unit root on the series to be modeled is essential. In order to perform this task, unit root tests, such as ADF, can be applied. One of key steps in this test procedure is to properly select the number of lags to be used. In this paper, we present in both quantitative and qualitative ways, that research on this subject is in expansion and there is no consensus on how to select the number of lags to be applied in the test. Therefore, it is evident that this is a subject in which further research is needed.

Key words: unit root, lag selection, electricity demand, stationarity

Introdução

Segundo a Agência Nacional de Energia Elétrica, ANEEL (2008), a energia é um fator de grande importância para o desenvolvimento econômico e social de um país, uma vez que fornece apoio mecânico, térmico e elétrico às ações humanas. Conforme EPE (2019), o fato de o consumo de energia estar relacionado à atividade econômica do país, faz com que a análise econômica seja relevante para compreender os movimentos de variação de consumo.

Uma vez que está diretamente relacionada ao crescimento econômico, como foi observado nas últimas décadas, sua produção adequada à demanda representa um fator relevante para a manutenção deste crescimento. No Brasil, existe a preocupação em elevar suas bases geradoras de energia, o que é evidenciado com a realização de fóruns nacionais, compostos por agentes ligados à questão energética, onde são discutidas medidas e metas para o setor (Viana & Silva, 2014).

Em De Assis Cabral, Legey e De Freitas Cabral (2017) é enfatizada a importância de se prever corretamente a demanda por eletricidade, onde é mostrado que as empresas geradoras devem garantir o fornecimento demandado com uma tolerância de 3%, para mais ou para menos, sob pena de serem multadas pela agência reguladora. Assim, estas empresas devem utilizar as melhores ferramentas disponíveis para realizar previsão da demanda de energia.

Portanto, o conhecimento prévio da evolução do consumo de energia elétrica pode, além de evitar problemas de abastecimento, servir de base para o planejamento do setor energético de um país e nortear políticas de eficiência energética, de modo que é fundamental que se estude a demanda de energia e os fatores que a influenciam. Para isso, é de suma importância buscar as melhores técnicas para que se realize a modelagem da demanda por eletricidade.

Dentre as maneiras mais utilizadas para analisar a demanda por energia elétrica, bem como realizar previsões, encontra-se a abordagem econométrica, com a análise de séries temporais, que pode ser definida como um conjunto de observações, ordenadas no tempo, de qualquer fenômeno aleatório (Bayer, 2010).



Nos últimos anos, trabalhos como os de Schmidt e Lima (2004), Irffi, Siqueira e Linhares (2009), Viana & Silva (2014) e De Oliveira, De Souza Tavares e Tavares (2018) utilizaram técnicas econométricas para realizar estudo da demanda por eletricidade no Brasil. No cenário mundial, alguns trabalhos recentes pode ser vistos em Al-Musaylh, Deo, Adamowski & Li (2018), Jamil (2020) e Murshed (2021), entre outros.

Em meio à análise econométrica, a presença de tendência estocástica (ou raiz unitária) na série, o que caracteriza a não estacionariedade, faz com que as previsões se tornem mais imprecisas conforme o distanciamento do último ponto da amostra aumente. Segundo Gujarati (2011), uma série é estacionária se sua média e variância são invariantes ao longo do tempo e o valor da covariância entre dois períodos depende somente do grau de defasagens entre as observações e não do período efetivo em que a covariância é calculada.

Deste modo, é importante identificar a presença de raiz unitária em uma série temporal a fim de encontrar a abordagem que melhor se aplica à característica da série estudada e, para isso, são utilizados, por exemplo, os testes de raiz unitária. Dentre estes testes, um dos mais populares, é o teste de Dickey Fuller (DF), proposto por Dickey e Fuller (1979), capaz de testar a presença de raiz unitária em modelos autorregressivos (AR). Cinco anos mais tarde, Said e Dickey (1984) propuseram uma modificação do teste DF, denominado teste DF aumentado (ADF), onde são incluídos valores defasados da diferença da série trabalhada. Com isso, o teste pode ser aplicado em séries que apresentem correlação nos resíduos.

Em ambos os testes supracitados, a hipótese nula indica a presença de raiz unitária e uma estatística τ é calculada, comparando-a com os valores tabelados por Said e Dickey (1984). Assim, de acordo com o nível de significância pré-estabelecido, toma-se a decisão de que há (ou não) indícios de presença de raiz unitária na série.

Para o teste ADF, existe a necessidade de selecionar o número de defasagens a serem aplicadas para realização do teste, entretanto, ainda não há, na literatura, um consenso sobre o procedimento a ser adotado. Desta maneira, cabe um maior aprofundamento nesta etapa de

aplicação do teste ADF, de modo que este trabalho busca, por meio de uma revisão sistemática, identificar lacunas científicas em relação a seleção do número de defasagens utilizado.

Procedimentos Metodológicos

Inicialmente se partirá de uma análise bibliométrica que, segundo Santos (2003), consiste em analisar quantitativamente a produção científica buscando encontrar indicadores cada vez mais confiáveis. Em seguida, será realizada a metassíntese qualitativa em que se buscará identificar lacunas e melhorias nas metodologias escolhidas na aplicação do teste ADF.

Para a definição do portfólio bibliográfico, foram escolhidas duas das maiores bases de dados de documentos científicos: *Web of Science* e *Scopus*. Na base *Web of Science*, a busca por documentos contendo as combinações de palavras-chave ocorreu por “Tópico”, enquanto na *Scopus*, a busca pelas palavras-chave ocorreu por “*Article title, Abstract, Keywords*”. A Tabela 1 apresenta as combinações de palavras, bem como o respectivo número de documentos encontrados.

Tabela 1

Resultados das combinações nas bases de dados

Palavras-chave	Web of Science	Scopus
<i>unit root x electricity demand</i>	60	21
<i>unit root x lag order</i>	37	24
<i>unit root x lag selection</i>	17	21
<i>unit root x adf x lag</i>	47	56
<i>unit root x adf x stationarity</i>	89	98

Fonte: Os autores.

Deste modo, foram encontrados 250 resultados na *Web of Science* e 220 resultados na *Scopus*, totalizando 470 documentos. Em seguida, foram retirados os documentos duplicados em cada base de dados, resultando em 231 na *Web of Science* e 196 na *Scopus*.

Na sequência, foram considerados apenas os documentos caracterizados como *articles* e *proceedings papers*, em cada base de dados, restando 229 resultados na *Web of Science* e 191

resultados na *Scopus*. Finalmente, foram selecionados apenas os artigos escritos nos idiomas inglês e espanhol, restando 225 na *Web of Science* e 185 na *Scopus*.

Posteriormente, foram cruzados os artigos das duas bases de dados com o auxílio da biblioteca *bibliometrix* (Aria e Cuccurullo, 2017) do *software R* e, com isso, foram eliminados os arquivos duplicados entre as bases de dados, obtendo um total de 305 artigos. A próxima etapa foi a leitura e filtragem dos 305 documentos a fim de identificar os que se enquadram diretamente ao tema da pesquisa. Após essa tarefa, permaneceram 29 arquivos, aproximadamente 10% do total dos que foram examinados na etapa.

Resultados e Discussões

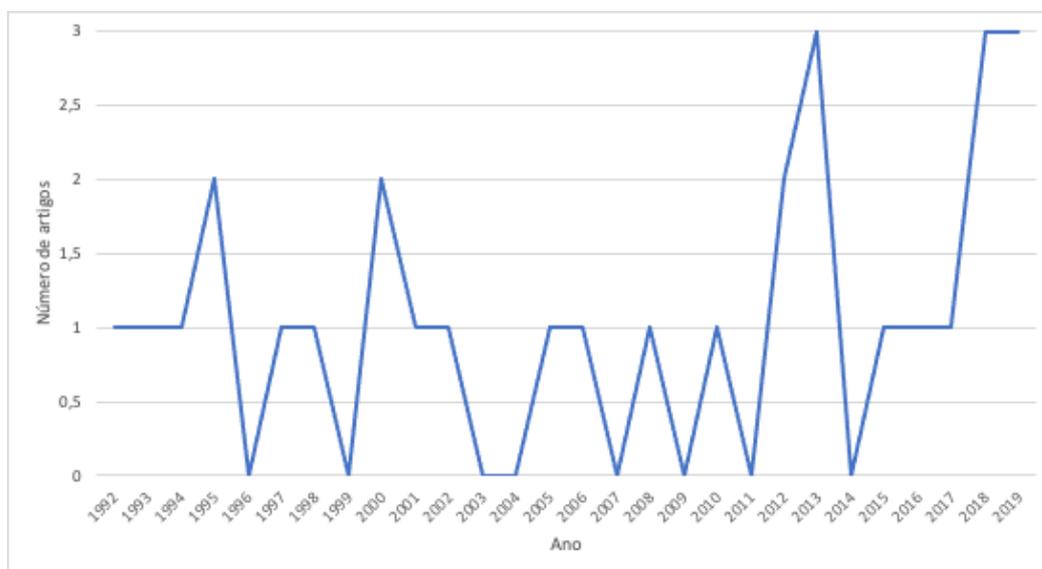
Análise Bibliométrica

De posse dos 28 artigos selecionados para a composição do portfólio bibliográfico final, foi dado início à análise bibliométrica.

A Figura 1 apresenta a distribuição dos artigos ao longo dos anos, onde é possível notar que, dentre os artigos selecionados, a ocorrência de três artigos em um mesmo ano se fez presente apenas em 2013, 2018 e 2019, evidenciando que o interesse pelo tema está crescendo.

Figura 1

Histórico de produção anual

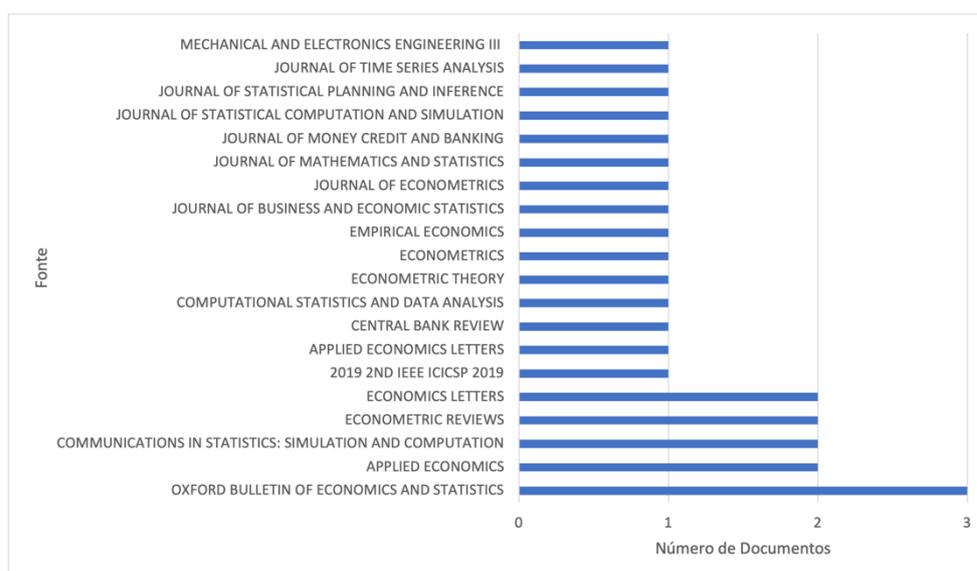


Fonte: Os autores.

De acordo com a Figura 2, dentre os periódicos de maior relevância, destaca-se o Oxford Bulletin of Economics and Statistics, com três publicações. Seguido de *Applied Economics*, *Communications in Statistics: Simulation and Computation*, *Economics Reviews* e *Economics Letters*, com duas publicações cada.

Figura 2

Publicações por Periódicos



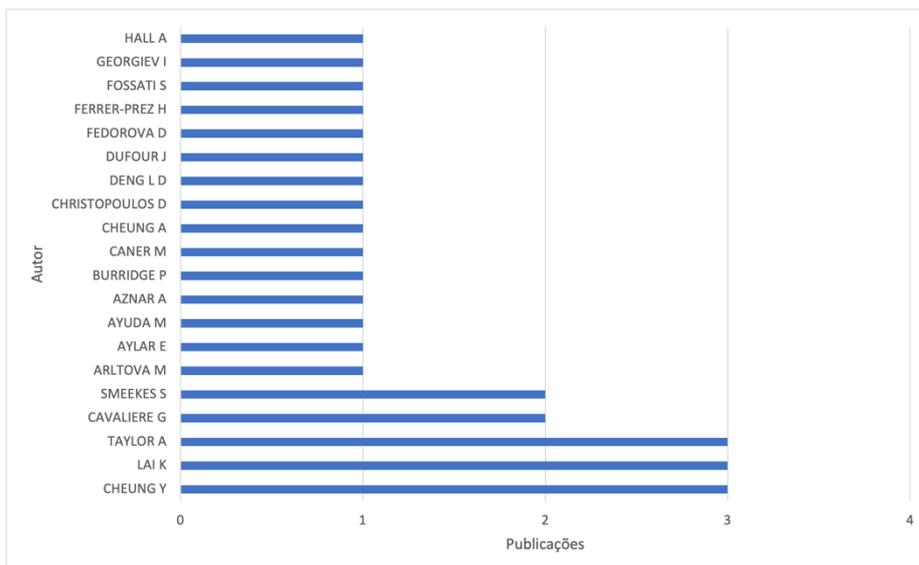
Fonte: Os autores.

Em relação aos autores, a Figura 3 mostra que os que mais publicaram são Cheung Y, Lai K, e Taylor A, com três publicações cada. Seguidos por Cavaliere G e Smeekes S, com duas cada um.



Figura 3

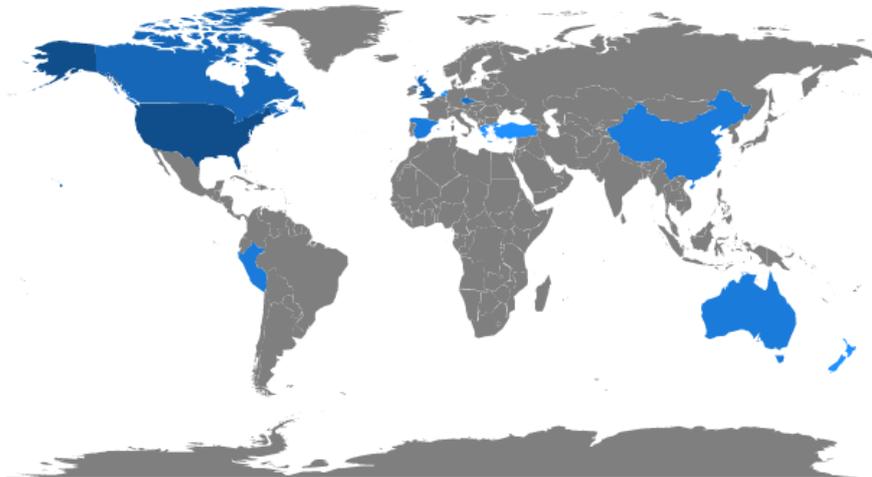
Publicações por autores



Fonte: Os autores.

Dentre os países com mais publicações sobre o tema, conforme a Figura 4, destacam-se Estados Unidos (um total de 10), seguido por Canadá e Reino Unido (quatro cada), e pela República Tcheca (três publicações). Cabe observar que estes quatro países representam 74,4% das publicações sobre o tema.

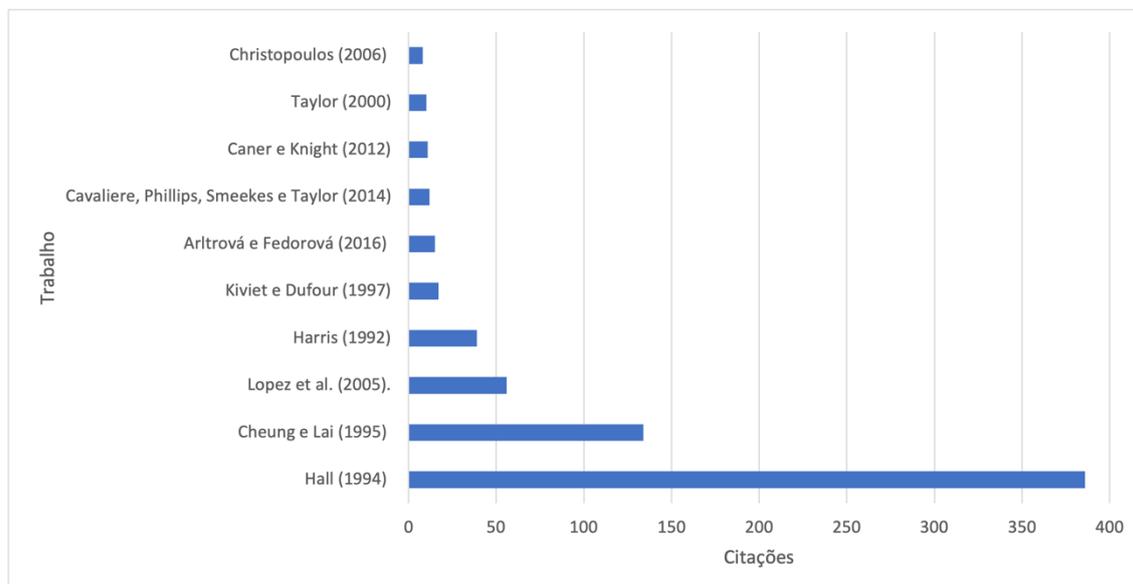
Figura 4

Publicações por país

Fonte: Os autores.

Os artigos mais citados são mostrados na Figura 5. Neste caso cabe observar que o artigo de Hall (1994) é o que apresenta maior número de citações: 386, cabendo menção a Cheung e Lai (1995) e Lopez et al. (2005), com mais 50 citações cada.

Figura 5

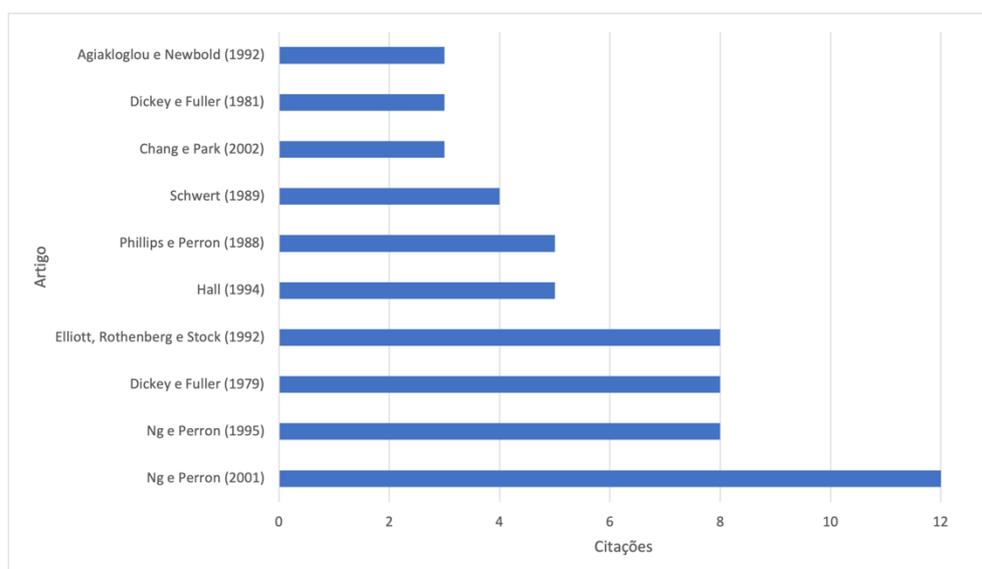
Citações por trabalhos

Fonte: Os autores

Em relação a análise das referências (Figura 6), o trabalho de Ng e Perron (2001) foi encontrado em doze artigos, seguido Ng e Perron (1995), Dickey e Fuller (1979) e Elliott, Rothenberg e Stock (1992) em oito, e logo por Hall (1994) e Phillips e Perron (1988), encontrados em cinco.

Figura 6

Referências mais citadas



Fonte: Os autores.

Os artigos selecionados são de 23 periódicos diferentes, tiveram em média 25,55 citações por ano, foram elaborados por 45 autores diferentes, sendo 12 destes em trabalho de único autor e 33 trabalhos multiautorais, resultando em 0,644 trabalho por autor.

Análise de conteúdo

Nesta etapa se busca encontrar lacunas científicas em relação ao tema estudado: seleção de defasagens na aplicação de testes de raiz unitária. Deste modo, foi construído o Quadro 1, onde constam os principais parâmetros relacionados a aplicação do teste ADF: comprimento da série, critério de seleção de defasagens, termos determinísticos e metodologia de desenvolvimento do trabalho.

Quadro 1

Principais parâmetros dos trabalhos que compõe o portfólio

Autores	Comprimento da Série (T)	Critério de seleção de defasagens	Termos determinísticos	Foco do trabalho
Aylar, Smeekes e Westerlund (2019)	50, 100, 200, 400, 800 e 1600	T^a (com sendo a variando de 0,05 a 0,49 em passos de 0,04).	Removidos através de procedimentos.	Simulação
Ferrer-Pérez, Ayuda e Aznar (2019)	100 e 250	DF_{OLS} , DF_{PQ} , EM, HL_{05} , HL_{06} , UR_{SA}	Presença de tendência determinística	Simulação
Zuo (2019)	-	-	-	Levantamento bibliográfico
Cavaliere, Georgiev e Taylor (2018).	100 e 500	$T^{1/3}$	Não são incluídos nas simulações.	Simulação
Burridge e Hristova (2017)	100, 250, 500 e 10000	AIC, BIC e HQ	Não são considerados	Demonstração matemática e simulação
Ozel e Ilalan (2017)	Dados diários em um período de 10 anos	-	Não são considerados	Aplicação
Quineche e Rodríguez (2017)	100 e 250	AIC, BIC, MAIC e MBIC	Presença de constante e tendência determinística	Simulação
Arltrová e Fedorová (2016)	25, 50, e 100	Critério de Schwert	Presença de constante e tendência determinística	Simulação
Paparoditis e Politis (2016)	Entre 50 e 1600	$T^{0,3}$	Não são incluídos nas simulações.	Simulação
Cavaliere, Phillips, Smeekes e Taylor (2014)	150 e 250	$\arg \min IC(k)$, $IC(k) = \ln \sigma_k^2 + k \frac{c_T}{T}$	Removidos através de procedimentos.	Simulação
Tam (2013)	50 e 100	Fixado, AIC, BIC e GS	Presença de constante e tendência determinística	Simulação
Fossati (2012)	100	AIC, BIC e MAIC	Presença de constante e tendência determinística	Simulação e Aplicação
Caner e Knight (2012)	100	AIC, BIC e MAIC	Presença de constante e tendência determinística	Simulação
Deng Lu (2011)	100	AIC, BIC, MAIC, MBIC e GS	Não são considerados	Simulação
Shaowen Wu (2010)	100, 150 e 250	GS e MAIC	Tendência removida	Simulação
Christopoulos (2006)	101	BIC	Presença de constante e	Aplicação

Autores	Comprimento da Série (T)	Critério de seleção de defasagens	Termos determinísticos	Foco do trabalho
			tendência determinística	
Lopez et al (2005)	104, 107, 117, 119 e 129	LM, GS e MAIC	Presença de constante e tendência determinística	Aplicação
Vougas (2002)	Entre 61 e 110	AIC	Presença de constante e tendência determinística	Aplicação
Weber (2001)	Entre 61 e 110	AIC, BIC, GS, SG e NAC	Presença de constante e tendência determinística	Aplicação
Taylor (2000)	50, 100 e 200	AIC, BIC e GS	Presença de constante e tendência determinística	Simulação
Montañés e Reyes (1999)	100	Fixado	Não são considerados	Simulação
Otero e Smith (1999)	1188	Fixado	Não são considerados	Simulação e Aplicação
Cheung e Lai (1998)	50, 100 e 150	AIC e BIC	Presença de constante e tendência determinística	Simulação
Kiviet e Dufour (1997)	500	Fixado	Intercepto	Simulação
Cheung e Lai (1995)	Entre 18 e 500	Valores entre 1 e 9	Presença de constante e tendência determinística	Simulação
Cheung e Lai (1995)	Entre 28 e 600	Fixado	Presença de constante e tendência determinística	Simulação
Em Hall (1994)	100 e 250	AIC, BIC e HQ	Presença de constante e tendência determinística	Simulação e Aplicação
Harris (1992)	50 e 100	Critério de Schwert	Presença de constante e tendência determinística	Simulação

Fonte: Os autores

Os trabalhos selecionados utilizaram séries temporais com amplitude variando entre 18 e 1.600 períodos, com dados reais ou obtidos por simulação, que foram caracterizados por modelos com diferentes estruturas. Além disso, utilizaram em torno de 24 critérios diferentes para

identificação da quantidade de lags. Além destas características principais, uma abordagem mais aprofundada sobre o tema é apresentada a seguir.

Metassíntese

Aylar, Smeekes e Westerlund (2019) mostram que aumentar o número de lags ótimo selecionados para o teste ADF está associado à diminuição do poder do teste. Os autores buscam avaliar a distribuição local assintótica do teste ADF quando os erros seguem um processo linear dirigido por inovações de diferença de Martingale - por meio de simulações -, o que pode fazer com que apresentem heterocedasticidade. Mostram que a influência do número de lags tende a diminuir conforme aumenta-se o número de elementos na série. Os autores afirmam que este teste é o mais popular entre os de raiz unitária, apresentando incontáveis aplicações, mas baixo poder.

Ferrer-Pérez, Ayuda e Aznar (2019) apresentam dois objetivos principais: identificar como a condição inicial afeta a seleção de lags e qual a implicação na performance do poder do teste. É assumida a presença de um outlier na condição inicial e são testadas séries com os erros considerados em três situações: ruído branco, processo autorregressivo e de médias móveis. Usam vários métodos para determinar a quantidade de lags, mas encontram que nenhum deles se apresenta superior para todas as circunstâncias, porém recomendam dois deles, deixando claro que o assunto ainda precisa de mais pesquisas.

Zuo (2019) compara as características das equações dos testes DF, ADF e outros dois comumente utilizados por meio do conteúdo de vinte materiais. Ele apresenta equações, características de testes e análise de hipóteses. Deste modo, o autor busca analisar equações, aplicações e distribuições dos testes por meio de levantamento bibliográfico.

Cavaliere, Georgiev e Taylor (2018) analisaram o teste ADF para o caso em que os choques seguem um processo linear conduzido por inovação de variância infinita, demonstrando matematicamente o que ocorre com a distribuição assintótica do teste nesta situação. Então é comparada a aplicação de Estimadores M com mínimos quadrados para a estimação do modelo. Logo, é aplicada uma técnica de reamostragem, por meio do algoritmo *Wild Bootstrap* (WB). Os



autores demonstram a validade do algoritmo WB para testes ADF, de modo que suas estatísticas alcançam os mesmos limites de distribuição que o teste ADF original, condicionadas à magnitude das inovações.

Burridge e Hristova (2017) estudam os estimadores de mínimos quadrados com a possibilidade de um processo não estacionário com no máximo uma raiz autorregressiva unitária. São realizados experimentos onde o número de defasagens é escolhido pelo BIC e pelo AIC. São utilizados modelos de ruído branco, bem como modelos AR, com o intuito de verificar se os critérios selecionam a ordem certa dos modelos. Os autores concluíram que uma escolha segura para aplicação dos testes seria realizar a seleção de defasagens por meio da equação $C(T) = \ln t(T)/T$, porém não deve ser adotada essa regra de maneira acriticamente.

Ozel e Ilalan (2017) propõem uma metodologia capaz de identificar a estacionariedade em séries de taxas de juros. Deste modo, convertem um processo OU (Ornstein–Uhlenbeck) em uma estrutura de regressão linear para calcular as estatísticas t pertencentes ao teste de reversão à média, e estes valores são comparados com os valores críticos obtidos por simulações de Monte Carlo. As séries estudadas são de taxas de juros diária correspondentes a dez anos da Turquia e dos Estados Unidos. São aplicados e comparados os testes OU, ADF e PP (Phillips–Perron), inclusive com a presença de tendência. O teste OU indicou estacionariedade nas séries, ao contrário do ADF e PP, para os casos em que todos os testes indicaram estacionariedade, o OU apresentou maiores valores- p . Entretanto, os autores indicam o teste apenas para os casos em que for necessário detectar reversão de média em séries de taxa de juros, e deve ser aplicado juntamente com o ADF para a detecção de reversão à média.

Quineche e Rodríguez (2017) apresentam um estudo baseado em simulação do tamanho e no comportamento excêntrico que pode ser apresentado pelo teste MGLS no contexto de mudanças estruturais. São gerados modelos AR(1) e MA(1) (Médias Móveis) com séries de tamanho 100 e 250 e considerando a presença de raiz unitária. São apresentados, para cada tamanho de série, os resultados do poder do teste para nove critérios diferentes de seleção de defasagens, além de seis

valores fixos de k . Os resultados mostraram uma grande relação entre valores explosivos negativos de MGLS e o valor da defasagem selecionado. Quando há quebra na série, os autores indicam a utilização do método supremo, selecionando as defasagens por MIC (Critério de Informação Modificado).

Arltrová e Fedorová (2016) têm por objetivo resolver o dilema de aconselhar qual teste de raiz unitária utilizar de acordo com critérios específicos. Para isso são apresentados os testes ADF, PP, ADF-GLS, NGP (Ng e Perron) e KPSS (Kwiatkowski, Phillips, Schmidt and Shin). São geradas séries AR(1), sem termos determinísticos, com o Φ variando de 0,01 a 0,99, com T sendo 25, 50, e 100 e realizadas 3.000 replicações com nível de significância de 5%. Em seguida, são aplicados e comparados os testes de raiz unitária através do poder do teste. Além disso, são realizadas simulações com $\Phi = 1$ para efeito de complementação. Para cada teste é mostrado o poder do teste em função do valor do parâmetro Φ para as séries dos diferentes tamanhos. A seguir é feita a comparação entre os testes para cada comprimento de série. Os autores formam uma tabela indicando a aplicação dos testes de acordo com os valores de Φ e com o comprimento da série. Ao fim, concluem que o trabalho fornece recomendações para a aplicação dos testes de raiz unitária de acordo com cada situação exposta.

Paparoditis e Politis (2016) apresentam uma nova maneira de determinar o número de lags a ser considerado no teste. Após realização de equacionamento matemático, é utilizada uma série real da variação de temperatura global anual e um estudo de simulação. Em seguida, é montado um experimento utilizando um modelo ARMA. É proposta uma fórmula para a determinação do número de lags: $k = T^a$, sendo testados diferentes valores de a , variando de 0,05 a 0,49. Ao fim, concluem que o valor mais apropriado para a é 0,3. Em seguida, apresentam prova matemática para sua conclusão. Os autores comparam sua proposta com o critério de informação de Akaike-AIC (Akaike, 1973), chegando à conclusão de que quando existe raiz unitária essa fórmula pode ser melhor, porém quando não há, é melhor usar o AIC e observam que a razão intuitiva para o reduzido poder



de teste do ADF é que como a ordem do modelo autorregressivo (p) tende ao infinito, os regressores de p se tornam assintoticamente colineares.

Cavaliere, Phillips, Smeekes e Taylor (2014) apontam que a grande parte dos trabalhos sobre o tema determinam o tamanho do lag em relação ao tamanho da série, porém poderia ser mais eficiente se fosse em relação às características dos dados. Os autores examinam o impacto de heterocedasticidade incondicional para amostras finitas, nos métodos que dependem dos dados para a seleção de lags. É proposto um novo método de seleção de lags que pode ser utilizado para heterocedasticidade incondicional. Deste modo, implementado um algoritmo WB (Wild Bootstrap). Os modelos testados são com $\phi = -c/T$, com $c=0$ e $c=7$. São realizadas simulações com diversos modelos variando os parâmetros de AR e MA, com erros homocedásticos e heterocedásticos. Por fim, é apresentado graficamente o poder do teste em função dos lags selecionados. Para os casos em que existe tendência, esta é removida através de algoritmo WB. O novo método de seleção de lags se mostrou capaz de evitar o problema de superestimar a ordem do lag, entregando testes com propriedades para amostras finitas quase idênticas aos correspondentes baseados nos métodos tradicionais, na presença de choques homocedásticos. A modificação proposta consiste em redimensionar os dados através de um estimador paramétrico do processo de volatilidade antes de computar a informação do critério de interesse. Os autores encontraram indícios de que, para diversos modelos, a seleção de lags tem sido altamente dependente do lag máximo permitido em certos casos. Enquanto isso, o critério proposto por eles não apresentou essa característica, enquanto na ausência de heterocedasticidade, apresentou resultados semelhantes aos testes tradicionais.

Tam (2013) apresenta uma análise de superfície resposta para o teste ADF com otimização da seleção da ordem do lag. As simulações mostram que para amostras finitas o resultado do teste depende diretamente da regra de seleção das defasagens. O teste apresenta distorção de tamanho quando a ordem do lag é otimizada, mas uma distribuição inapropriada é empregada a implementação do teste. São propostas superfícies resposta para distribuições de amostras finitas



com otimização da seleção de lags. São realizados dez experimentos para cada para cada cruzamento de T com de k_{\max} para a realização dos experimentos. Então são geradas séries com tamanho $T+100$ e descartadas as primeiras 100 observações a fim de descartar o efeito do valor inicial. O autor conclui que o comportamento do teste ADF para amostras finitas é sensível ao método de seleção de lags aplicado.

Fossati (2012) analisa a seleção de lags para testes de raiz unitária covariáveis. Para isso, são gerados pares de séries com diferentes características para cada par. Em seguida, são aplicados os testes EJ (Elliott e Jansson), e CADF-GLS (estimação por mínimos quadrados generalizados do teste DF com inclusão de covariáveis). Para cada um dos testes são utilizados diferentes critérios de seleção de defasagens e os resultados comparados entre si. Em seguida, é utilizada a série de inflação dos EUA, utilizando como covariáveis o PIB e a taxa de desemprego. O autor mostra que o AIC e o critério de informação bayesiano (BIC) selecionam lags muito pequenos que resultam em testes com grandes distorções. Um critério de informação modificado pode ser utilizado para construir um teste de raiz unitária covariável com bom tamanho e poder. São aplicados os critérios modificados e BIC em um conjunto de dados que contém a inflação nos EUA combinada com outras variáveis econômicas, onde segundo o critério modificado, a hipótese nula não é rejeitada, encontrando resposta diferente do BIC.

Caner e Knight (2012) apresenta um modo alternativo de identificar a raiz unitária utilizando os chamados estimadores Bridge. A seleção de parâmetros se dá por meio do MBIC (critério de informação bayesiano modificado). Em seguida, são comparados os resultados dos testes DF-GLS com Bridge e Bridge com ADF, para casos com e sem tendência. Não existe grande ênfase na seleção de defasagens, entretanto são comparados os valores críticos considerando k_{\max} como 24 e 36. É mostrado que os estimadores Bridge podem selecionar o modelo correto com probabilidade tendendo a um. Por fim, o trabalho utiliza o estimador Bridge para determinar os parâmetros que são zero e diferentes de zero, assim o modelo diferencia a presença ou não de raiz unitária e seleciona o lag ótimo simultaneamente.



Deng Lu (2011) simula modelos modelo autorregressivo integrado de médias móveis - ARIMA(0,1,q) e compara a influência de diferentes métodos de seleção de defasagens e o poder do teste ADF. São geradas séries ARIMA(0,1,q) e aplicado o teste ADF utilizando os critérios AIC, BIC, MAIC (AIC modificado), MBIC e GS (geral para específico). Em seguida são simulados diferentes valores de θ , com T e k_{\max} fixados em 100 e 20, respectivamente, com 10.000 simulações para cada caso, e os resultados apresentam o lag médio selecionado por cada método, bem como o percentual de rejeição da hipótese nula em cada um. O artigo mostra que os critérios de GS, AIC e MAIC tendem a selecionar mais lags do que os BIC e MBIC. Além disso, o lag selecionado aumenta quando o valor da raiz tende à raiz unitária. Ao final conclui que o critério de informação modificado apresentou tamanho mais apropriado, enquanto o critério geral para específico mostrou mais robustez nas propriedades do teste ADF.

Shaowen Wu (2010) aborda a temática da seleção de lags no teste DF-GLS, que é uma modificação do ADF onde os dados têm sua tendência removida. O artigo reinvestiga o problema de seleção de lags no teste, mostrando a importância de utilizar valores críticos ajustados para o tamanho da amostra na seleção de lags para a aplicação do teste. Para isso, são estimados os modelos de DF-GLS e é utilizado k_{\max} como $\text{int}\left[12\left(\frac{T}{100}\right)^{\frac{1}{4}}\right]$. Os níveis de significância trabalhados são 5 e 10%. São testados os critérios de informação modificados de Akaike estimados por OLS (mínimos quadrados ordinários) e GLS (mínimos quadrados generalizados), bem como testes sequenciais t . Foram selecionadas as defasagens através de cada teste e cada método para taxa de câmbio de 20 países. Assim, é mostrado o número de defasagens selecionado bem como o resultado de cada teste. Em seguida são gerados um modelo MA e um AR e é avaliado o tamanho do teste em cada situação. Em seguida é gerado um novo modelo, com mais semelhanças às séries de taxas de juros a fim de avaliar o tamanho do teste nesta situação para cada um dos testes. No caso em que os dados imitam a estrutura de séries de câmbio, a seleção de lags através de testes sequenciais t , para o teste DF-GLS considerando o nível de significância de 10% apresentou o melhor resultado. Por outro lado, quando a série apresenta grandes coeficientes negativos MA, a seleção por critério de informação

apresentou-se vantajoso. Pelo fato de nenhum dos critérios ter sido superior para diferentes tipos de dados, o autor sugere a utilização deles em conjunto.

Christopoulos (2006) utiliza modelos não lineares para investigar a estacionariedade do PIB per capita dos sete países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), apresentando o modelo de TAR (autoregressivos com threshold) que une as hipóteses de não linearidade e estacionariedade. Também é aplicado o teste ADF, utilizando critério de Schwert (1989) e testes-t sequenciais para o PIB de sete países. Para os mesmos países, membros da OCDE, são aplicados os testes TAR e estimado o modelo de Fourier. Em seguida são comparados os resultados dos testes ADF e TAR, concluindo que os testes não lineares se aplicam melhor às séries de PIB.

Lopez et al. (2005) estudam a aplicação de testes de raiz unitária em séries de Paridade de Poder de Compra (PPC), dando enfoque na seleção ótima de defasagens no resultado dos testes. São utilizadas séries anuais de 16 países com tamanhos diferentes para a aplicação dos testes ADF e DF-GLS e os critérios de seleção de defasagens de Akaike Modificado, testes-t sequenciais e LM (teste do Multiplicador de Lagrange). São apresentados os resultados dos testes e o número de defasagens selecionado para as situações com e sem tendência. Em seguida, é elaborada uma tabela com o número de rejeições para cada teste para fins de comparação. Por fim, os autores aplicam os testes escolhendo o número de defasagens que maximiza o valor p do teste de raiz unitária. O resultado do artigo apresentou maior taxa de rejeição da hipótese nula tanto o ADF e quanto para o DF-GLS, se comparados ao estudo anterior. Os autores concluem que o número de defasagens pode afetar o resultado dos testes em séries de PPC.

Vougas (2002) utiliza as séries de Nelson e Plosser (NP) para aplicação do teste DF. O autor afirma que em NP foi escolhido o número de defasagens de maneira incorreta e que havia presença de heterocedasticidade condicional e que o número de lags é selecionado através de comparação da autocorrelação da série em nível e em primeira diferença, porém uma abordagem mais correta seria a utilização de critérios de informação, como, por exemplo, o AIC. Assim, são utilizadas as séries de NP e como método de seleção de defasagens o AIC, e para aliviar o efeito de heterocedasticidade são



construídas matrizes de covariância heterocedasticidade consistente-t. São construídas tabelas com os resultados e comparações entre eles. O autor encontrou resultados diferentes de NP em quatro das quatorze séries e conclui que a não rejeição da hipótese nula deve ser tratada com cuidado, pois pode estar relacionada a erro de especificação no teste.

Weber (2001) aborda a seleção de defasagens para os testes de raiz unitária das séries de NP, no caso em que a hipótese alternativa inclui uma quebra de tendência. O autor diz que diferentes escolhas do número de defasagens podem afetar o poder do teste ADF de maneira significativa. Para as séries abordadas, a detecção de raiz unitária pode ser altamente sensível ao método de seleção de lags. São comparados cinco diferentes critérios de seleção de defasagens. O autor conclui que para as séries macroeconômicas estudadas, quando não é o critério geral para específico a ser utilizado, a evidência contra a presença de raiz unitária em sete das séries enfraquece. Além disso, em vários casos a evidência contra a raiz unitária enfraquece consideravelmente. Quando a quebra é determinada endogenamente, como em Zivot e Andrews, os resultados obtidos pelo critério alternativo de seleção de lags fornece ainda menos evidência contra a presença de raiz unitária para as séries de NP. Por fim, o autor sugere que ao trabalhar com séries macroeconômicas os pesquisadores devem checar cuidadosamente o número de defasagens selecionado para testar a presença de raiz unitária.

Taylor (2000) afirma que, para amostras finitas, o teste ADF é altamente sensível às duas formas de inclusão de variáveis determinísticas e à estrutura de defasagens adotada. Em relação à seleção por meio de testes sequenciais, é demonstrado que este efeito aparentemente é atribuível ao viés sistemático de tendência a zero dos lags estimados, que é acentuado na presença de termos determinísticos. Para os experimentos, são geradas séries ARIMA(p,1,q), com 10.000 replicações e são considerados de pares de k_{\max} e k_{\min} de (5,0), (10,0) e (10,4). As séries são geradas com e sem termos determinísticos, e são utilizados três critérios de seleção de defasagens: AIC, BIC e GS, considerando o nível de significância de 5%. Os autores encontram evidências de que para os modelos baseados nos dados com termos determinísticos tendem a subestimar o número de

defasagens nos testes ADF, deste modo a rotina de inclusão de uma tendência linear na regressão do ADF não é aconselhada. Uma solução para este problema poderia ser utilizar métodos de seleção que tendem a selecionar menos lags. Outra hipótese seria aumentar o nível de significância para a tomada de decisão, como por exemplo para 0,20 e aplicar os testes sequenciais. Por fim, outra sugestão seria estimar o modelo sem os termos determinísticos e utilizar algum algoritmo baseado nos dados para estimar o valor de k , depois reestimar a regressão com os termos determinísticos.

Montañés e Reyes (1999) examinam o comportamento assintótico dos testes ADF e PP na presença de quebras estruturais. Quando a mudança na tendência da série é omitida, a probabilidade de rejeição da hipótese nula é leve, porém essa característica só é observada para amostras finitas, não assintoticamente. Os autores replicam 10.000 vezes cada experimento para os testes ADF e PP considerando $k = 1, 2, 3, 4, 5, 6$ e 12, são abordados os casos com perturbação positiva, sem perturbação e com perturbação negativa. Em seguida é elaborada uma tabela com o poder do teste para cada situação e cada valor de k . É mostrado que a magnitude da quebra afeta as duas estatísticas fazendo com que sejam deslocadas para a zona de aceitação, conforme aumenta a magnitude da quebra. Em relação ao número de defasagens para o teste ADF, valores altos induzem o pesquisador a aceitar a presença de raiz unitária, mesmo para quebras de baixa magnitude.

Otero e Smith (1999) estudam os efeitos de aumentar a frequência de observação e da extensão dos dados no teste de integração de Johansen (λ_{\max}). Para isso são criadas duas séries com 1.188 observações mensais. São testados os valores de poder do teste para λ_{\max} considerando séries mensais, trimestrais ou anuais. Em seguida, os testes são aplicados em dados reais de variação de taxa de juros. Os autores concluem que é preferível ter dados coletados por um longo período do que ter muitas observações em um período curto.

Cheung e Lai (1998) examina o problema de como seleção de defasagens usando critério de informação padrão pode ajudar a melhorar o poder de teste do ADF. São construídas curvas empíricas do poder de teste sob diferentes métodos de seleção de defasagens. O estudo reavalia o impacto de critérios de informação baseado nos dados no teste ADF pela correção para o viés para



amostras finitas. Para isso, são simulados modelos AR e MA com k variando de 1 a 10, Φ de 0,7 a 1 com passo de 0,025 e são variados os coeficientes de AR e MA. O número máximo de defasagens é calculado por $4(T/100)^{1/4}$ e são aplicados testes com e sem tendência. São replicadas 40.000 vezes cada experimento, sendo níveis de significância adotados de 5 e 10%. Para os modelos AR e MA, são apresentados graficamente os resultados através do poder do teste em relação ao valor de Φ para valores fixados diferentes de k e valores selecionados por AIC, com e sem tendência. Ao final, concluem que para os casos em que as séries apresentam comportamento AR os critérios padrões são úteis, porém na presença de erros MA os resultados não são encorajadores, uma vez que a performance neste caso não se apresentou superior à seleção via um número fixo defasagens.

Kiviet e Dufour (1997) implementam testes de raiz unitária em modelos ARX (autorregressivo com entradas exógenas) de regressão dinâmica, os testes incluem a seleção do lag na presença de quebras estruturais, para séries com diferentes características. O estudo indica que os testes apresentam especificação adequada apenas assintoticamente e são não exatos para amostras finitas. É desenvolvido um processo alternativo de inferência para modelos linear de regressão dinâmica e demonstrada sua viabilidade através de modelos empíricos, sendo denominados testes exatos. Por fim, os autores apresentam suas próprias estatísticas de testes obtidas através de simulações de Monte Carlo. Então concluem que os testes padrões podem apresentar p-valores menores do que os exatos, porém isso não implicaria em maior poder de teste. Se a distribuição dos erros não apresentar normalidade, a precisão dos modelos assintóticos é muito pobre para amostras finitas, enquanto dos testes exatos apresentou melhor análise sensível a vários tipos de distribuições.

Cheung e Lai (1995) têm por objetivo demonstrar o efeito da seleção de defasagens nos valores críticos do teste DF-GLS para amostras finitas, além de estimar os valores críticos corretos considerando o efeito da ordem de defasagem. Para isso são utilizadas equações de superfícies respostas e realizadas 296 combinações de T e k , considerando valores críticos a 5% e 25.000 replicações. São calculados os valores críticos e colocados em uma tabela com cada valor de defasagem. Na continuidade são aplicados os testes com os valores críticos estimados em séries reais

de PIB per capita e taxas de câmbio. Os autores concluem que a seleção de defasagens pode afetar o resultado do teste ADF-GLS para o caso de amostras finitas.

Cheung e Lai (1995) afirmam que embora a estatística de teste do ADF não dependa da ordem do lag assintoticamente, para amostras finitas, a estatística pode ser sensível ao número de lags. Então os autores abordam casos de amostras finitas com diferentes tamanhos para determinar a ordem do lag na aplicação do teste. São realizados 30.000 experimentos para valores de T menores do que 30, e 40.000 quando T é maior do que 30. Com isso, são construídas superfícies respostas através de uma equação para a estimação do valor crítico, de k e de T. Por fim, medem a velocidade com que os valores de amostra finita conseguem se aproximar de seus valores assintóticos, concluindo que para os casos de constante e tendência a aproximação ocorre mais lentamente. O estudo mostra que a ordem do lag, além do tamanho da amostra, pode afetar o comportamento do teste ADF para amostras finitas.

Hall (1994) considera o impacto de pré-testes baseados na seleção do modelo nos processos de seleção de defasagem do teste ADF. São utilizados critérios de informação e GS, além disso também são atribuídos valores fixos para k . São gerados modelos AR e MA e são calculados valores de níveis de significância de acordo com o valor de defasagem selecionado ou critério adotado. Os resultados são apresentados para T=100 e 250, com os valores de Φ sendo 0,9, 0,95 e 1,00. Por fim, os testes são aplicados em dados reais de inventário. O autor conclui que para o caso de modelos ARIMA(p,1,0) seria mais satisfatório utilizar GS, HQC, AIC ou BIC. Para os modelos AR e que nenhum dos valores é nulo, é recomendado utilizar HQC ou BIC. Ainda, é mais prudente estimar k por vários critérios diferentes antes de tomar uma decisão sobre a presença de raiz unitária.

Harris (1992) testa a influência de aspectos operacionais no resultado do teste ADF, especialmente o número de defasagens, considerando modelos AR, ARMA e ARIMA. O autor explica o procedimento do teste ADF e apresenta suas propostas para a realização de experimentos. São geradas T+20 observações e são descartadas as 20 iniciais e cada modelo é replicado 2.000 vezes. Para cada modelo são apresentados os resultados considerando a presença de constante e



tendência, somente constante, e nenhuma delas. O número de defasagens parte do critério de Schwert e é selecionado por meio do critério de Akaike. Em outro experimento, o número de defasagens é escolhido de acordo com o valor, entre 0 e 25, que maximiza o R^2 . Para análise dos resultados, é apresentado para as séries de diferentes comprimentos e características, o número de defasagem selecionado, o valor da estatística de teste e a estatística t padrão. Os autores consideram o nível de significância de 5% para tomar as decisões. Os resultados mostram que selecionar as defasagens através da maximização de R^2 se mostrou insatisfatório, pois o número de defasagens selecionado apresentou alto desvio padrão. O tamanho correto do teste foi consistentemente alcançado quando o AIC foi utilizado para escolher o número ideal de defasagens.

Conclusões

Este artigo foi proposto para construir um portfólio bibliográfico relacionado à seleção de defasagens em testes de raiz unitária. Para isso, foram identificados, por meio de análise bibliométrica, os principais autores, periódicos e países em que está sendo abordado o tema.

Em relação à análise de conteúdo, a metassíntese mostrou que o tema de seleção de defasagens na aplicação dos testes de raiz unitária vêm sendo tema de muitos estudos nos últimos anos. Grande parte destes estudos utiliza o teste ADF para identificar a presença de raiz unitária, evidenciando sua popularidade. Entretanto, diversos trabalhos propõem novos testes e estes geralmente não são capazes de superar o ADF nos resultados, salvo em algumas situações em específico. É possível observar frequentes relatos de baixo poder do teste ADF, bem como uma possível sensibilidade do resultado do teste tanto com as defasagens selecionadas, quanto com a inclusão de termos determinísticos.

Deste modo, a proposta deste trabalho foi alcançada, visto que foi possível constatar que o tema vem sendo objeto de estudo nos anos recentes em diversos países. Além disso, cabe salientar que a seleção de defasagens em testes de raiz unitária se encontrou como fator relevante no resultado do teste, podendo ser tema para aprofundamento do estudo deste tipo de testes. Para

trabalhos futuros, é possível adicionar novas palavras chaves a fim de testar diferentes combinações de pesquisas nas bases de dados.

Referências

- Al-Musaylh, M. S., Deo, R. C., Adamowski, J. F., & Li, Y. (2018). Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. *Advanced Engineering Informatics*, 35, 1-16.
<https://doi.org/10.1016/j.aei.2017.11.002>
- Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of informetrics*, 11(4), 959-975.
- ANEEL, A. (2008). Atlas de energia elétrica do Brasil. *Brasília*.
- Aylar, E., Smeekes, S., & Westerlund, J. (2019). Lag truncation and the local asymptotic distribution of the ADF test for a unit root. *Statistical Papers*, 60(6), 2109-2118.
<https://doi.org/10.1007/s00362-017-0911-y>
- Bayer, F. M. (2010). Wavelets e modelos tradicionais de previsão: um estudo comparativo. *Rev. Bras. Biom*, 28(2), 40-61.
- Cavaliere, G., Georgiev, I., & Taylor, A. R. (2018). Unit root inference for non-stationary linear processes driven by infinite variance innovations. *Econometric Theory*, 34(2), 302-348.
<https://doi.org/10.1017/S0266466616000037>
- Burrige, P., & Hristova, D. (2008). Consistent estimation and order selection for nonstationary autoregressive processes with stable innovations. *Journal of Time Series Analysis*, 29(4), 695-718. <https://doi.org/10.1111/j.1467-9892.2008.00579.x>
- Cavaliere, G., Phillips, P. C., Smeekes, S., & Taylor, A. R. (2014). Lag length selection for unit root tests in the presence of nonstationary volatility. *Econometric Reviews*, 34(4), 512-536.
- Cheung, Y. W., & Lai, K. S. (1995). Lag order and critical values of the augmented Dickey–Fuller test. *Journal of Business & Economic Statistics*, 13(3), 277-280.
<https://doi.org/10.1080/07350015.1995.10524601>



- Cheung, Y. W., & Lai, K. S. (1998). Power of the augmented dickey-fuller test with information-based lag selection. *Journal of Statistical Computation and Simulation*, 60(1), 57-65.
<https://doi.org/10.1080/00949659808811871>
- de Assis Cabral, J., Legey, L. F. L., & de Freitas Cabral, M. V. (2017). Electricity consumption forecasting in Brazil: A spatial econometrics approach. *Energy*, 126, 124-131.
<https://doi.org/10.1016/j.energy.2017.03.005>
- de Oliveira, V. H., de Souza Tavares, R., & Tavares, L. A. (2018). Demanda residencial por energia elétrica no Brasil (2004-2015). *Revista Catarinense de Economia*, 2(1), 142-162.
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), 427-431.
<https://doi.org/10.1080/01621459.1979.10482531>
- Elliott, G., Rothenberg, T. J., & Stock, J. H. (1992). Efficient tests for an autoregressive unit root.
<https://doi.org/10.3386/t0130>
- EPE. (2019). Atlas da eficiência energética no Brasil em 2019 /Empresa de Pesquisa Energética- Rio de Janeiro. Disponível em: <http://epe.gov.br/pt/publicacoes-dadosabertos/publicacoes/atlas-da-eficiencia-energeticabrasil-2019>.
- Fedorová, D. (2016). Selection of unit root test on the basis of length of the time series and value of ar (1) parameter. *Statistika*, 96(3), 3.
- Ferrer-Pérez, H., Ayuda, M. I., & Aznar, A. (2019). The sensitivity of unit root tests to the initial condition and to the lag length selection: A Monte Carlo Simulation Study. *Communications in Statistics-Simulation and Computation*, 50(4), 1062-1072.
<https://doi.org/10.1080/03610918.2019.1577967>
- Gujarati, D. N., & Porter, D. C. (2011). *Econometria básica-5*. Amgh Editora.
- Hall, A. (1994). Testing for a unit root in time series with pretest data-based model selection. *Journal of Business & Economic Statistics*, 12(4), 461-470.

Harris, R. I. (1992). Testing for unit roots using the augmented Dickey-Fuller test: Some issues relating to the size, power and the lag structure of the test. *Economics letters*, 38(4), 381-386.

[https://doi.org/10.1016/0165-1765\(92\)90022-Q](https://doi.org/10.1016/0165-1765(92)90022-Q)

Irffi, G., Castelar, I., Siqueira, M. L., & Linhares, F. C. (2009). Previsão da demanda por energia elétrica para classes de consumo na região Nordeste, usando OLS dinâmico e mudança de regime. *Economia Aplicada*, 13, 69-98.

Jamil, R. (2020). Hydroelectricity consumption forecast for Pakistan using ARIMA modeling and supply-demand analysis for the year 2030. *Renewable Energy*, 154, 1-10.

<https://doi.org/10.1016/j.renene.2020.02.117>

Lopez, C., Murray, C. J., & Papell, D. H. (2005). State of the art unit root tests and purchasing power parity. *Journal of Money, Credit and Banking*, 361-369.

Murshed, M. (2021). Modeling primary energy and electricity demands in Bangladesh: An Autoregressive distributed lag approach. *Sustainable Production and Consumption*, 27, 698-712. <https://doi.org/10.1016/j.spc.2021.01.035>

Ng, S., & Perron, P. (1995). Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, 90(429), 268-281.

Ng, S., & Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica*, 69(6), 1519-1554. <https://doi.org/10.1111/1468-0262.00256>

Otero, J., & Smith, J. (2000). Testing for cointegration: power versus frequency of observation—further Monte Carlo results. *Economics Letters*, 67(1), 5-9. [https://doi.org/10.1016/S0165-1765\(99\)00245-1](https://doi.org/10.1016/S0165-1765(99)00245-1)

Papadoditis, E., & Politis, D. N. (2016). The asymptotic size and power of the augmented Dickey-Fuller test for a unit root. *Econometric Reviews*, 37(9), 955-973.

<https://doi.org/10.1080/00927872.2016.1178887>

Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2),

335-346. Disponível em: <https://doi.org/10.1093/biomet/75.2.335>

- Quineche, R., & Rodríguez, G. (2017). Selecting the lag length for the MGLS unit root tests with structural change: A warning note for Practitioners based on simulations. *Econometrics*, 5(2), 17. <https://doi.org/10.3390/econometrics5020017>
- Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3), 599-607. <https://doi.org/10.1093/biomet/71.3.599>
- Santos, R. N. M. D. (2003). Produção científica: por que medir? O que medir?. *Revista digital de Biblioteconomia e Ciência da Informação*, 1(1).
- Schmidt, C. A. J., & Lima, M. A. (2004). A demanda por energia elétrica no Brasil. *Revista brasileira de economia*, 58, 68-98. <https://doi.org/10.1590/S0034-71402004000100004>
- Schwert, G. W. (2002). Tests for unit roots: A Monte Carlo investigation. *Journal of Business & Economic Statistics*, 20(1), 5-17. <https://doi.org/10.1198/073500102753410354>
- Viana, G. I. M. N., & Silva, A. L. M. (2014). Um modelo para projeções para demanda por energia elétrica, 2009-2017 para o setor residencial no Brasil. *Revista Brasileira de Energia*, 20, 107-126.
- Zuo, X. (2019, September). Several Important Unit Root Tests. In 2019 IEEE 2nd International Conference on Information Communication and Signal Processing (ICICSP) (pp. 10-14). IEEE. <https://doi.org/10.1109/ICICSP48821.2019.8958557>