

FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

DANIEL FERREIRA ARRUDA

**PREVENDO O CONSUMO AGREGADO DAS FAMÍLIAS A PARTIR DE MODELOS
VEC, VAR E REGRESSÕES MIDAS**

SÃO PAULO
2023

DANIEL FERREIRA ARRUDA

**PREVENDO O CONSUMO AGREGADO DAS FAMÍLIAS A PARTIR DE MODELOS
VEC, VAR E REGRESSÕES MIDAS**

Dissertação apresentada à Escola de
Economia de São Paulo da Fundação
Getulio Vargas, como requisito para
obtenção do título de Mestre em
Economia

Área de concentração: Economia

Orientador: Prof. Dr. Marcel Ribeiro

SÃO PAULO

2023

Arruda, Daniel Ferreira.

Prevendo o consumo agregado das famílias a partir de modelos VEC, VAR e regressões MIDAS / Daniel Ferreira Arruda. - 2023.

40 f.

Orientador: Marcel Bertini Ribeiro.

Dissertação (mestrado profissional MPE) – Fundação Getulio Vargas, Escola de Economia de São Paulo.

1. Consumo (Economia) - Brasil. 2. Família - Aspectos econômicos. 3. Modelos econométricos. I. Ribeiro, Marcel Bertini. II. Dissertação (mestrado profissional MPE) – Escola de Economia de São Paulo. III. Fundação Getulio Vargas. IV. Título.

CDU 330.567.2(81)

DANIEL FERREIRA ARRUDA

**PREVENDO O CONSUMO AGREGADO DAS FAMÍLIAS A PARTIR DE MODELOS
VEC, VAR E REGRESSÕES MIDAS**

Dissertação apresentada à Escola de Economia de São Paulo da Fundação Getúlio Vargas, como requisito para obtenção do título de Mestre em Economia

Área de concentração: Economia

Data de aprovação: 01 / 03 / 2023

Banca Examinadora:

Prof. Dr. Marcel Ribeiro (Orientador)
FGV-EESP

Prof. Dr. Emerson Marçal
FGV-EESP

Prof. Dr. Fábio Gomes
USP-FEA RP

RESUMO

Esse trabalho tem por objetivo avaliar a previsibilidade da série brasileira de consumo agregado das famílias em diferentes horizontes de tempo. Para o curto e médio prazo, são estimados modelos VEC e VAR que capturam a dinâmica entre o consumo agregado das famílias e seus principais determinantes. Para o curtíssimo prazo (*nowcasting*), são estimadas regressões *Mixed Data Sampling* (MIDAS) empregando as vendas no varejo ampliado, a confiança do consumidor e o IBC-Br, e variando o número de meses conhecidos dos regressores dentro do trimestre previsto em cada especificação. Os resultados apontam que os modelos de correção de erro trazem leve melhora do desempenho preditivo em relação ao passeio aleatório, especialmente quando incorporam a variação do crédito às famílias ou o serviço da dívida. Além disso, as regressões MIDAS mostram boa capacidade de prever o consumo com regressores conhecidos dentro do trimestre, superando novamente o passeio aleatório, e é possível afirmar que o MIDAS restrito possui desempenho superior ao irrestrito (U-MIDAS).

Palavras-chave: Consumo agregado; Previsão; Correção de erros; *Nowcasting*; MIDAS; U-MIDAS

ABSTRACT

This paper aims to evaluate the predictability of the Brazilian household aggregate consumption data in different time horizons. For the short and medium term, VEC and VAR models are estimated to capture the dynamics between aggregate household consumption and its main determinants. For the very short term (nowcasting), *Mixed Data Sampling* (MIDAS) regressions are estimated using amplified retail sales, consumer confidence and the IBC-Br, and varying the number of months known to the regressors within the forecasted quarter in each specification. Results indicate that the error correction models bring a slight improvement in the predictive performance in relation to the random walk, especially when including household credit growth or debt service. In addition, the MIDAS regressions show a good capability to predict consumption with regressors known within the quarter, again outperforming the random walk, and it is possible to state that the restricted MIDAS performs better than the unrestricted version (U-MIDAS).

Keywords: Aggregate consumption; Forecasting; Error correction; Nowcasting; MIDAS; U-MIDAS

LISTA DE FIGURAS

Figura 1 – Variáveis utilizadas nos modelos de curto e médio prazo.....	15
Figura 2 – Variáveis utilizadas nos modelos de curtíssimo prazo	17
Figura 3 – Coeficientes dos modelos MIDAS e U-MIDAS.....	28

LISTA DE TABELAS

Tabela 1 – Correlação entre o consumo e a defasagem de cada determinante (em logaritmo)	16
Tabela 2 – Correlação entre o consumo e os determinantes (em logaritmo)	18
Tabela 3 – Acurácia das previsões fora da amostra para modelos de curto e médio prazo	25
Tabela 4 – Acurácia das previsões fora da amostra para modelos de curtíssimo prazo	29
Tabela 5 – Teste ADF para variáveis dos modelos de curto e médio prazo	36
Tabela 6 – Procedimento de Johansen para modelos de curto e médio prazo.....	36
Tabela 7 – Estimativas para modelos de curto e médio prazo	37
Tabela 8 – P-valores do MCS para cada horizonte de previsão do curto e médio prazo	37
Tabela 9 – Teste ADF para variáveis dos modelos de curtíssimo prazo.....	38
Tabela 10 – Coeficientes dos modelos MIDAS	39
Tabela 11 – Coeficientes estimados para os modelos U-MIDAS.....	40
Tabela 12 – Parâmetros estimados para os modelos MIDAS	40
Tabela 13 – P-valores do MCS para o curtíssimo prazo	40

SUMÁRIO

1 INTRODUÇÃO	8
2 REVISÃO DA LITERATURA	10
3 METODOLOGIA	14
3.1 DADOS.....	14
3.1.1 Curto e médio prazo	14
3.1.2 Curtíssimo prazo	16
3.2 MODELOS ECONOMETRICOS	18
3.2.1 Correção de erros	18
3.2.2 Regressão MIDAS.....	20
4 RESULTADOS.....	23
4.1 CURTO E MÉDIO PRAZO	23
4.2 CURTÍSSIMO PRAZO.....	27
5 CONCLUSÃO	31
REFERÊNCIAS.....	33
APÊNDICES	36

1 INTRODUÇÃO

De acordo com o IBGE (2022), o consumo das famílias é o maior componente do PIB brasileiro pela ótica da demanda, tendo correspondido a 61,0% deste indicador no ano de 2021. Dessa forma, a compreensão dos fatores econômicos determinantes do consumo é fundamental para qualquer análise prospectiva a respeito da evolução do nível da atividade doméstica e do bem-estar da população.

Em particular, a capacidade de antever a resposta do consumo agregado a partir de diferentes conjunturas econômicas é do interesse de firmas, analistas de mercado e formuladores de políticas públicas. Por consequência, a modelagem empírica e a proposição de um arcabouço de projeção adequado para os dados brasileiros têm grande importância para diversos agentes na economia.

Em um horizonte de curto e médio prazo, parte da literatura econômica foca na utilização de modelos de correção de erro para a modelagem conjunta do consumo agregado e seus determinantes, como sugerido em Campbell (1987), uma vez que a possível convergência para o nível de equilíbrio implica em certa previsibilidade para essas variáveis alguns períodos a frente. Ademais, alguns trabalhos, principalmente na literatura internacional, avaliam a contribuição de variáveis macroeconômicas, como o crédito e o endividamento do setor privado, para a previsibilidade do consumo agregado.

Já sob uma ótica de curtíssimo prazo, a disponibilidade de indicadores macroeconômicos de frequência mensal levanta a possibilidade de uso desses dados para realizar o *nowcasting* do consumo agregado, cuja divulgação geralmente ocorre cerca de dois meses depois do fim do período de referência. Além disso, metodologias econométricas mais recentes, como as regressões *Mixed Data Sampling* (MIDAS) propostas por Ghysels, Santa-Clara e Valkanov (2004) e Ghysels, Sinko e Valkanov (2007), permitem o uso de regressores de frequência mais alta, em um modelo de projeção de variável de frequência mais baixa, sem a necessidade de agregação temporal prévia das variáveis explicativas, o que minimiza a perda de informação preditiva potencialmente útil.

Essa dissertação fornece evidências de leve redução da raiz do erro quadrático médio (RMSE, em inglês) fora da amostra, em relação a dois *benchmarks*

(passeio aleatório e AR(1)), ao utilizar modelos de correção de erro para projetar o consumo agregado em diferentes horizontes de tempo. Em particular, há ganhos de previsão estatisticamente significantes para o horizonte de um trimestre à frente (ao nível de 5,0%). Ademais, os modelos de correção de erros que incorporam o crédito às famílias e o serviço da dívida mostram ganhos de acurácia preditiva em relação às demais especificações estimadas.

A relevância dessas variáveis traz algumas conclusões para a condução da política econômica no curto prazo. Por um lado, os resultados sugerem que o consumo das famílias apresenta certa sensibilidade a choques macroeconômicos que impactem a disponibilidade de crédito ou o comprometimento da renda com dívidas. Por outro lado, as evidências apresentadas nessa dissertação também sugerem maior eficácia da política econômica em influenciar o nível de atividade no curto prazo, com destaque para medidas que estimulem a concessão de crédito ou limitem a pressão do endividamento excessivo sobre o balanço das famílias.

Já os modelos MIDAS, em geral, apresentam queda expressiva e estatisticamente significativa do RMSE para o *nowcasting* do consumo agregado utilizando variáveis de frequência mensal, na comparação aos mesmos *benchmarks* univariados. Vale destacar, ainda, que a especificação restrita do MIDAS tem melhor acurácia preditiva em relação à irrestrita. Esses resultados, portanto, validam o uso dessa metodologia para realizar um diagnóstico tempestivo da evolução da atividade econômica, contribuindo para uma melhor tomada de decisão de agentes de mercado e formuladores de políticas públicas.

O restante dessa dissertação está dividido em quatro partes adicionais. A segunda parte sintetiza alguns trabalhos nacionais e internacionais associados à explicação e previsão do consumo agregado. A terceira parte descreve os dados utilizados e explica brevemente as metodologias de modelos de correção de erro e de regressões MIDAS. A quarta parte apresenta as estimativas realizadas e explica os resultados mais importantes obtidos. Por fim, a quinta parte conclui a dissertação e sugere uma possível extensão a esse trabalho.

2 REVISÃO DA LITERATURA

Em um dos primeiros trabalhos associados à explicação do consumo agregado, Hall (1978) mostra que, sob a Hipótese do Ciclo de Vida – Renda Permanente (HCV-RP), e supondo utilidade quadrática e juros fixos, o consumo deve seguir um passeio aleatório com tendência. Nesse caso, a única variável com poder preditivo sobre o consumo agregado seria sua primeira defasagem, que supostamente já incorporaria as informações de todos os demais determinantes econômicos, como a renda.

Posteriormente, Campbell (1987) demonstra que, sob a HCV-RP, a combinação linear entre consumo e renda seria estacionária, sugerindo a aplicação de modelos de correção de erro para explicar essas variáveis. Em particular, o autor encontra evidências de que o aumento da poupança ajuda a prever quedas na renda do trabalho. Essa relação decorre da hipótese de que os consumidores, ao anteciparem uma redução futura da renda, limitariam o consumo contemporâneo e aumentariam a poupança.

Na sequência, diversos trabalhos empregam modelos de correção de erros para explicar a relação entre o consumo agregado e seus determinantes (AL-EYD et al., 2006; DE BONDT; GIESECK; ZEKAITE, 2020; FERNANDEZ-CORUGEDO; PRICE; BLAKE, 2003; LETTAU; LUDVIGSON, 2001; LUDVIGSON; STEINDEL, 1998). Em geral, os autores utilizam variáveis de renda e riqueza – e, eventualmente, outros regressores (como juros e endividamento) – e obtêm evidências heterogêneas com respeito à capacidade desse tipo de modelo de explicar o consumo agregado.

Para além das variáveis de renda e riqueza, já comumente empregadas na literatura, o crédito às famílias surge como uma possível variável adicional para auxiliar na previsão do consumo, como sugerem artigos focados no teste da equação de Euler. Por um lado, Ludvigson (1999) apresenta um arcabouço teórico que justifica a correlação entre o crescimento previsto do crédito e do consumo agregado, observada empiricamente. Por outro, Brady (2008) fornece evidências de que a relevância do crédito para explicar o consumo se reduziu ao longo do tempo, apontando para uma maior aderência da HCV-RP em anos recentes, nos Estados Unidos.

Embora o crédito possa ter impacto positivo sobre a demanda, o endividamento excessivo e a subsequente necessidade de desalavancagem pode comprometer a renda dos consumidores e gerar incentivos para redução futura do consumo. De fato, os trabalhos de Estrada et al. (2014), Kim, Setterfield e Mei (2015) e Albuquerque e Krustev (2018) corroboram esse argumento e fornecem evidências de impacto negativo de medidas de endividamento sobre o consumo das famílias, sugerindo que essa variável pode ser útil na estimação de modelos de previsão.

Na literatura nacional, os primeiros trabalhos buscaram avaliar a adequação das séries brasileiras aos resultados teóricos esperados com base na HCV-RP. Vale destacar as conclusões obtidas por Gomes (2010), segundo as quais o consumo agregado no Brasil depende em larga medida da renda, o que pode decorrer tanto de restrição ao crédito quanto de miopia, mas mostra baixa sensibilidade à taxa de juros. Por sua vez, Lopes (2017) encontra evidências de que o arcabouço de correção de erros é adequado para modelar a dependência do consumo em relação à renda. Embora esses trabalhos, em geral, não foquem na previsão do consumo, os resultados obtidos são úteis para a definição das variáveis e metodologias que possam ser empregadas para a proposição de um modelo de projeção para o caso brasileiro.

Ainda na literatura nacional, Schettini et al. (2012) fornecem evidências de menor erro quadrático médio, na projeção do consumo agregado fora da amostra, ao utilizar técnicas de cointegração e séries de renda disponível e crédito às famílias, na comparação à outras especificações estimadas. Por fim, Felini, Gomes e Soave (2022) sugerem que a confiança do consumidor e o retorno do mercado de ações também são preditores robustos para explicar o consumo na análise dentro da amostra, partindo de um conjunto de diversas variáveis e especificações distintas.

Dentro da literatura nacional, a primeira parte dessa dissertação se aproxima mais do realizado em Schettini et al. (2012), na medida em que é avaliada a previsibilidade do consumo agregado fora da amostra a partir de técnicas de cointegração. No entanto, o presente trabalho se diferencia por incluir novos determinantes do consumo, como o índice do mercado de ações (*proxy* da riqueza financeira) e o comprometimento de renda com o serviço da dívida, e por avaliar a

adequação das projeções obtidas para diferentes horizontes de tempo e em um período mais recente.

Em geral, o consumo agregado é calculado em frequência trimestral, enquanto alguns regressores estão disponíveis em frequência mais alta. Embora a agregação prévia do dado de alta frequência permita contornar esse problema, essa solução descarta informações potencialmente úteis para a estimação de modelos preditivos. Uma resposta mais rigorosa a esse problema é dada por Ghysels, Santa-Clara e Valkanov (2004) e Ghysels, Sinko e Valkanov (2007), que propõem os modelos *Mixed Data Sampling* (MIDAS). Em síntese, essa metodologia permite incorporar em um modelo toda a informação disponível em regressores de alta frequência, ao mesmo tempo em que limita a proliferação dos coeficientes estimados.

Internacionalmente, os modelos MIDAS têm sido bastante empregados na previsão de variáveis econômicas. Vale citar o trabalho de Armesto, Engemann e Owyang (2010), no qual os autores apontam para queda do RMSE na projeção do PIB ao utilizar o MIDAS com o regressor (emprego) conhecido dentro do trimestre previsto, na comparação a outras formas de agregação temporal mais simples. Ademais, Foroni e Marcellino (2014) mostram que o MIDAS tem bom desempenho na previsão do PIB e seus componentes na Zona do Euro, mesmo na comparação às outras abordagens de frequência mista, como as equações *bridge* e o MF-VAR.

Com foco no consumo agregado, Dreger e Kholodilin (2013) encontram ganhos preditivos ao utilizar a metodologia MIDAS e diferentes reponderações dos componentes da confiança do consumidor como regressores. Com objetivo análogo, Duarte, Rodrigues e Rua (2017) fornecem evidências otimistas sobre a capacidade preditiva do MIDAS empregando dados coletados de *Automated Teller Machines* (ATM) e *Points-of-Sale* (POS), bem como séries de vendas no varejo e confiança do consumidor.

Na literatura nacional, Zuanazzi e Ziegelmann (2014) utilizam diversos regressores econômicos e financeiros de frequência mensal com o objetivo de projetar o PIB brasileiro, e apontam que tanto o MIDAS restrito quanto o irrestrito permitem redução do erro de projeção fora da amostra, em relação ao modelo ARMA. Da mesma forma, Ferreira (2020) se vale da metodologia MIDAS para

projetar o crescimento real do PIB no Brasil. Ao contrário de Zuanazzi e Ziegelmann (2014), a autora foca em incorporar regressores macroeconômicos para explicar a evolução da atividade. As conclusões desse trabalho apontam para a superioridade do MIDAS restrito em relação ao irrestrito e para a importância de incluir o IBC-Br em modelos de previsão da atividade econômica.

3 METODOLOGIA

3.1 DADOS

Essa seção descreve as bases de dados utilizadas nessa dissertação, e está subdividida entre as séries empregadas nos modelos de curto e médio prazo e nos modelos de curtíssimo prazo.

3.1.1 Curto e médio prazo

O modelo base estimado nessa parte do trabalho usa as séries temporais do consumo agregado, do Produto Interno Bruto (PIB) e do índice da bolsa de valores (*proxy* da riqueza financeira). As duas primeiras séries são extraídas das Contas Nacionais Trimestrais do IBGE, para o período entre o primeiro trimestre de 1996 e o segundo trimestre de 2022. Já para representar a riqueza financeira, é utilizado o índice *ibovespa*, construído em frequência trimestral a partir dos retornos mensais extraídos do IPEADATA.

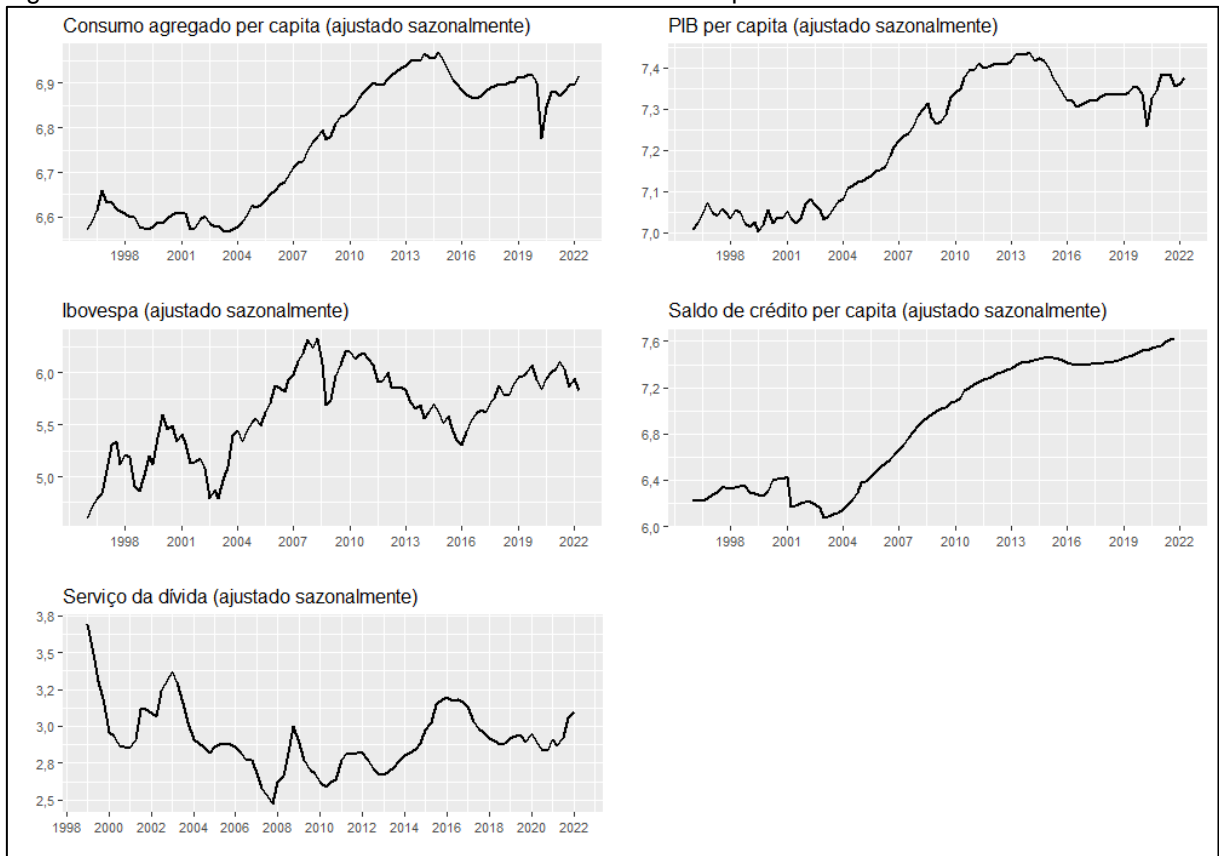
O modelo alternativo 1 inclui também uma variável do estoque de crédito às famílias. Com o objetivo de obter uma série temporal mais longa, é utilizado o dado disponibilizado pelo *Bank for International Settlements* (BIS), disponível desde o terceiro trimestre de 1994, como feito em Felini, Gomes e Soave (2022). Vale ressaltar que essa série apresenta um *outlier* relevante no segundo trimestre de 2001, quando se observa uma queda abrupta do saldo de crédito às famílias. De acordo com Banco Central do Brasil (2001), uma possível explicação para essa variação é a cessão de parte relevante da carteira de financiamento imobiliário da Caixa Econômica Federal para outra empresa, constituindo um fator exógeno que deve ser controlado na estimação dos modelos.

Por fim, o modelo alternativo 2 adiciona uma variável de serviço da dívida ao conjunto de dados do modelo base. Essa série, também calculada pelo BIS, representa a razão entre o pagamento de juros e amortizações em relação à renda do setor privado não financeiro, e está disponível apenas desde o primeiro trimestre de 1999, o que restringe o tamanho da amostra usada na estimação do modelo que incorpora essa variável.

Para o deflacionamento das variáveis (exceto do serviço da dívida como fração da renda), é utilizado o deflator do consumo agregado, calculado pela divisão

da série nominal pela real. Além disso, todas as séries são ajustadas sazonalmente, por meio da aplicação da metodologia do *X13 ARIMA-SEATS*, implementada no pacote *seasonal* do *software* estatístico R (SAX; EDDELBUETTEL, 2018). Por fim, o PIB, o consumo agregado e o estoque de crédito às famílias são transformados em termos per capita, a partir do uso da série de população residente do IBGE¹. As séries transformadas são mostradas em logaritmo natural na Figura 1.

Figura 1 – Variáveis utilizadas nos modelos de curto e médio prazo



Fontes: IBGE, IPEA, BIS. Elaboração própria.

De forma a obter uma análise preliminar acerca das correlações entre as variáveis utilizadas no trabalho, a Tabela 1 apresenta o coeficiente angular (β) e o coeficiente de determinação (R^2) de uma regressão simples entre a primeira diferença do logaritmo do consumo e a primeira diferença do logaritmo de cada determinante com uma defasagem (incluindo o próprio consumo). A única exceção é o serviço da dívida, que é incluído em nível, por ser estacionário (como se verifica na sequência do trabalho). Ademais, uma vez que a pandemia de coronavírus teve forte

¹ Essa série é extraída do IPEADATA em base anual e interpolada com o uso da função *Spline* do *software* estatístico R.

impacto sobre as séries utilizadas, as estimativas são mostradas para a amostra total e para a amostra até 2019.

Tabela 1 – Correlação entre o consumo e a defasagem de cada determinante (em logaritmo)

Amostra total	Δ Consumo agregado	Δ PIB	Δ lbovespa	Δ Saldo de crédito	Serviço da dívida
β	-0,0009	0,1301	0,0362*	0,1526**	-0,0191*
R ²	0,00%	1,61%	5,36%	8,22%	4,44%
Amostra até 2019	Δ Consumo agregado	Δ PIB	Δ lbovespa	Δ Saldo de crédito	Serviço da dívida
β	0,2798**	0,2843***	0,0274**	0,1562***	-0,0182***
R ²	7,94%	12,24%	7,39%	21,78%	12,65%

Elaboração própria. *, ** e *** indicam significância a 5%, 1% e 0,1%, respectivamente.

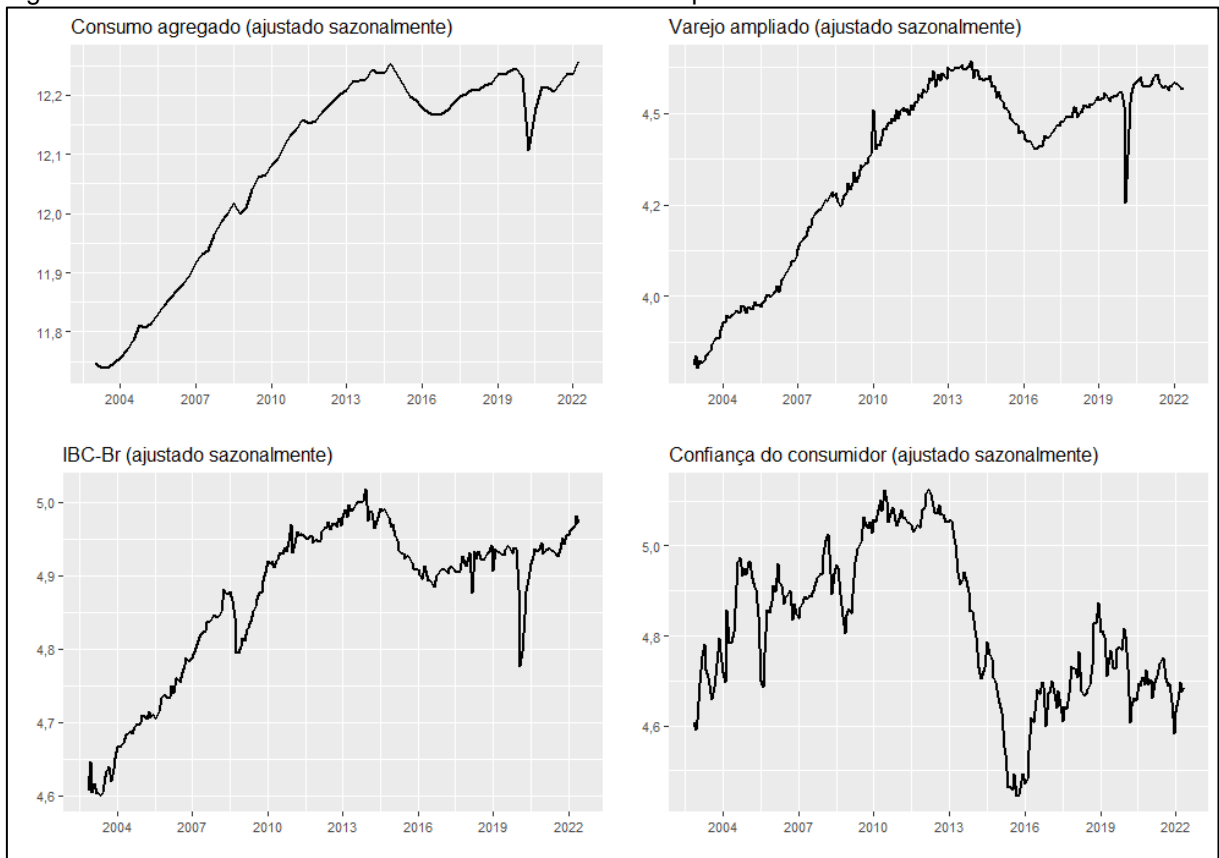
Os resultados mostram que as correlações entre as séries são mais evidentes ao se expurgar o efeito distorcivo da pandemia de coronavírus, como se atesta pelo aumento do coeficiente de determinação para todas as variáveis na amostra até 2019. Além disso, na amostra restrita, vale destacar que todos os determinantes individualmente mostram correlação de sinal esperado com o consumo agregado em um trimestre à frente.

3.1.2 Curtíssimo prazo

Para os determinantes dos modelos de curtíssimo prazo, parte-se, inicialmente, das séries mensais de vendas no varejo e da confiança do consumidor, em linha com o realizado em Duarte, Rodrigues e Rua (2017). Adicionalmente, opta-se por incluir o IBC-Br como variável explicativa, em função dos resultados otimistas encontrados para essa série na explicação do PIB, como evidenciado por Ferreira (2020). Vale ressaltar que, nessa etapa do trabalho, o consumo agregado não é tomado em termos per capita, como feito para os modelos de curto e médio prazo.

Para as vendas no varejo, é utilizado o índice de volume de vendas no comércio varejista ampliado, calculado e deflacionado pelo IBGE, e disponibilizado para o período a partir de janeiro de 2003. O IBC-Br, que representa uma *proxy* mensal do PIB, é calculado e deflacionado pelo Banco Central, e está disponível para o mesmo período das vendas no varejo ampliado. Por último, de forma a obter uma série mais longa, é utilizado o índice de confiança do consumidor calculado pela FecomercioSP, seguindo a escolha feita por Felini, Gomes e Soave (2022). As séries são ajustadas sazonalmente por meio do *X13 ARIMA-SEATS* e mostradas em logaritmo natural na Figura 2.

Figura 2 – Variáveis utilizadas nos modelos de curtíssimo prazo



Fontes: IBGE, Banco Central, FecomercioSP. Elaboração própria.

Assim como feito na seção anterior, a Tabela 2 traz uma análise preliminar acerca da correlação entre a diferença do logaritmo do consumo agregado e a diferença do logaritmo dos determinantes descritos acima, limitando a amostra até 2019 para excluir os efeitos da pandemia de coronavírus. Para contornar o problema da diferença de frequência entre as variáveis, os dados mensais são previamente agregados por meio da média simples dentro de cada trimestre. Além disso, dado que um dos objetivos desse trabalho é utilizar as variáveis de alta frequência para realizar o *nowcasting* do consumo agregado, o coeficiente angular (β) e o coeficiente de determinação (R^2) são reportados tanto a partir de uma regressão com os dados em termos contemporâneos, quanto com a variável explicativa defasada em um trimestre.

Tabela 2 – Correlação entre o consumo e os determinantes (em logaritmo)

Série defasada	Δ Varejo ampliado	Δ IBC-Br	Δ Confiança do cons.
β	0,3091***	0,2973**	0,0341
R ²	31,49%	13,47%	3,27%
Série contemporânea	Δ Varejo Ampliado	Δ IBC-Br	Δ Confiança do cons.
β	0,4013***	0,5365***	0,0509*
R ²	51,28%	42,33%	7,05%

Elaboração própria. *, ** e *** indicam significância a 5%, 1% e 0,1%, respectivamente.

A primeira conclusão extraída da tabela acima é de que pode haver um ganho significativo em utilizar variáveis contemporâneas para projetar o consumo agregado, em especial no caso do varejo ampliado e do IBC-Br, embora a confiança do consumidor também se mostre relevante. Além disso, o sinal do coeficiente angular é positivo em todos os casos, conforme esperado. Ademais, vale destacar que as variáveis agregadas em termos defasados também possuem certa capacidade de explicar a variação corrente do consumo agregado, embora menor do que se usadas em termos contemporâneos. O uso da metodologia MIDAS, justamente por não requerer a agregação prévia dos regressores, será importante para avaliar se há informação útil contida nos dados mensais para além das leituras contemporâneas.

3.2 MODELOS ECONOMETRÍCOS

Essa seção define e explica as especificações econométricas utilizadas nessa dissertação, com foco particular nos modelos de correção de erro e nas regressões MIDAS.

3.2.1 Correção de erros

De acordo com Engle e Granger (1987), ainda que algumas séries temporais possam ter tendências estocásticas individualmente, é possível que exista uma combinação linear entre essas variáveis que seja estacionária. Com base nisso, os autores definem que um conjunto de séries é cointegrado de ordem (d,b) se cada componente é integrado de ordem d e existe ao menos uma combinação linear entre as séries, imposta pelo vetor de cointegração, de ordem (d-b).

Para fins preditivos, a importância desse conceito reside no fato de que, se há cointegração, as variáveis de um sistema tenderão a retornar ao equilíbrio dado pela relação linear de longo prazo. Dessa forma, faz sentido modelar um conjunto de variáveis cointegradas a partir de um VAR ampliado por um termo de correção de

erro (resultando em um modelo VEC), cujo vetor de cointegração e os coeficientes de ajustamento de cada variável estão associados a uma matriz estimada π . Esse modelo é explicitado abaixo, onde X representa um vetor de p variáveis, conforme Johansen (1991).

$$\Delta X_t = \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Pi X_{t-k} + \Phi D_t + \mu + \varepsilon_t \quad (1)$$

Nessa dissertação, a existência de cointegração entre as séries temporais analisadas é testada por meio do procedimento de Johansen (JOHANSEN, 1991, 1995)². O número de defasagens utilizadas em cada modelo é selecionado com base no critério de informação BIC do VAR em nível e em testes de autocorrelação dos resíduos.

A aderência dos modelos estimados é avaliada, principalmente, por meio do erro de projeção fora da amostra. Para isso, os modelos são obtidos, inicialmente, até o último trimestre de 2013, deixando o período entre 2014 e 2019 (24 observações) para avaliação dos resultados. Os dados após 2020 são excluídos da análise em função da distorção gerada pela pandemia de coronavírus sobre as relações entre as variáveis, como discutido na análise preliminar. As projeções são realizadas para um, dois, quatro e seis passos à frente, com expansão gradual da amostra de estimação.

Posteriormente, as previsões obtidas são comparadas às de um modelo de passeio aleatório e de um AR(1), por meio da razão do RMSE. Adicionalmente, é utilizado o teste proposto por Diebold e Mariano (1995), com modificações de Harvey, Leybourne e Newbold (1997), doravante denominado DM, para avaliar se há um ganho estatisticamente significativo dos modelos propostos em relação a cada um dos dois *benchmarks*. Por fim, busca-se determinar o conjunto que contém os modelos estatisticamente superiores, dado um nível de confiança, a partir da obtenção do *Model Confidence Set* (MCS), desenvolvido por Hansen, Lunde e Nason (2011). Para isso, utiliza-se o pacote *MCS* do *software* estatístico R (BERNARDI; CATANIA, 2018).

² Caso haja cointegração, o modelo VEC é estimado por meio do método de máxima verossimilhança, utilizando o pacote *TsDyn* do *software* estatístico R (STIGLER, 2020); caso não haja cointegração, é estimado o modelo VAR das variáveis em diferenças a partir do pacote *vars* do mesmo *software* (PFAFF, 2008).

3.2.2 Regressão MIDAS

A metodologia MIDAS foi introduzida em Ghysels, Santa-Clara e Valkanov (2004) e Ghysels, Sinko e Valkanov (2007) para lidar com a modelagem de variáveis de frequências distintas. O modelo proposto pelos autores pode ser sintetizado nas equações abaixo.

$$y_t = \beta_0 + \beta_1 B(L^{1/m}; \theta) x_t^{(m)} + \varepsilon_t^{(m)} \quad (2)$$

$$B(L^{1/m}; \theta) = \sum_{k=0}^K B(k; \theta) L^{k/m} \quad (3)$$

$$L^{1/m} x_t^{(m)} = x_{t-1/m}^{(m)} \quad (4)$$

Nesse modelo, o operador $L^{1/m}$ simboliza as defasagens da variável de alta frequência, enquanto o polinômio $B(k; \theta)$, que pode assumir diferentes formatos, é empregado para gerar os coeficientes para cada defasagem do regressor por meio da estimação de um número reduzido de parâmetros θ . Considerando uma variável y trimestral e x mensal, o coeficiente m terá valor três, uma vez que, para cada observação de y , haverá três observações contemporâneas de x .

A partir da especificação acima, é possível inferir que o número de coeficientes estimados em um modelo MIDAS irrestrito pode crescer significativamente na medida em que aumenta o número de defasagens utilizadas da variável explicativa³. Desse modo, a imposição de uma forma funcional para $B(k; \theta)$ é importante para limitar o número de parâmetros, ao mesmo tempo em que confere certa flexibilidade para os coeficientes obtidos. Nessa dissertação, opta-se por utilizar a função exponencial de Almon para modelar $B(k; \theta)$, baseada na especificação proposta em Ghysels, Sinko e Valkanov (2007). Essa função, mostrada abaixo, é interessante pois permite obter diferentes formatos para a curva dos coeficientes em relação às defasagens empregadas no modelo, com base na estimação de poucos parâmetros.

³ Supondo y trimestral e x mensal, e que os dados dos últimos três trimestres do regressor (inclusive o corrente) impactem a previsão de y , o modelo sem restrição polinomial implica na estimação de nove coeficientes (3x3) apenas para a variável x .

$$B(k; \theta) = \frac{e^{\theta_1 k + \theta_2 k^2}}{\sum_{k=1}^K e^{\theta_1 k + \theta_2 k^2}} \quad (5)$$

Em particular, vale ressaltar que, quando θ_1 e θ_2 são iguais a zero, a função implica em pesos iguais para todas as defasagens, enquanto uma queda dos pesos é garantida na medida em que o segundo coeficiente seja negativo. No entanto, vale citar que essa restrição impõe que o modelo seja estimado por mínimos quadrados não lineares (MQNL). Todos os resultados de modelos MIDAS nesse trabalho são obtidos através do pacote *midasr* do *software* estatístico R (GHYSELS; KVEDARAS; ZEMLYS, 2016).

Apesar do claro apelo ao uso de uma função polinomial que limite o número de coeficientes estimados, alguns trabalhos apontam para superioridade preditiva de modelos MIDAS irrestritos (U-MIDAS) em relação a *benchmarks* univariados, como Duarte, Rodrigues e Rua (2017). Essa especificação consiste, simplesmente, na estimação de um coeficiente para cada defasagem de cada variável explicativa, que pode ser feita por Mínimos Quadrados Ordinários (MQO). Uma vez que, nessa dissertação, há uma diferença de frequência pequena entre as variáveis explicativas e explicada (três dados mensais para cada trimestral), e apenas três regressores, o risco de proliferação de coeficientes é relativamente baixo, o que permite comparar empiricamente a adequação dos modelos U-MIDAS em relação ao MIDAS restrito tradicional.

A adequação dos modelos estimados é avaliada por meio de projeções fora da amostra, a partir de três especificações distintas, como em Dreger e Kholodilin (2013): a primeira inclui os regressores até o primeiro mês do trimestre da variável explicada; a segunda considera os regressores conhecidos até o segundo mês desse trimestre; a terceira pressupõe que todos os dados mensais já são conhecidos no instante de tempo anterior à divulgação do consumo agregado. Conforme feito anteriormente, a estimação inicial leva em conta os dados até o final de 2013, deixando os 24 trimestres subsequentes para avaliação das projeções. A seleção de defasagens é feita a partir do critério de informação BIC para o caso do MIDAS irrestrito, com limite superior de oito; para o MIDAS restrito, a defasagem máxima é fixada em oito, independente do critério de informação, uma vez que o

número de coeficientes estimados não depende do número de defasagens utilizadas.

As projeções obtidas em todas as especificações são comparadas às de um modelo de passeio aleatório e de um AR(1), com base no cálculo da razão do RMSE. Além disso, é avaliado se as previsões dos modelos estimados são estatisticamente distintas desses dois *benchmarks* por meio do teste de DM. Por fim, é empregada novamente a metodologia do MCS, de forma a obter o conjunto que contém os modelos estatisticamente superiores ao nível de confiança desejado.

4 RESULTADOS

4.1 CURTO E MÉDIO PRAZO

O primeiro passo para a avaliação de cointegração entre as variáveis utilizadas no trabalho é a aplicação do teste ADF (DICKKEY; FULLER, 1979) para definir a ordem de integração de cada série individualmente, tomando as amostras restritas até o último trimestre de 2013. A Tabela 5, no Apêndice A, mostra que as variáveis são integradas de ordem um, a exceção do serviço da dívida, que é integrado de ordem zero.

A partir daí, é avaliada a existência de cointegração para todos os modelos, por meio do procedimento de Johansen, como mostrado na Tabela 6 do Apêndice A. A escolha do número de defasagens foi feita partindo do sugerido pelo critério de informação BIC para o VAR em nível e acrescentando defasagens gradativamente caso o modelo anterior apontasse para autocorrelação dos resíduos. Para o modelo alternativo 1, foi incluída uma *dummy* para o tratamento do *outlier* da série de crédito descrito anteriormente. Para todos os modelos, o procedimento é realizado em duas especificações distintas, sendo que ambas incorporam uma constante na relação de cointegração, mas apenas a segunda inclui uma constante nas equações das variáveis em diferenças.

Para o modelo base, os resultados do procedimento de Johansen são mistos, apontando para zero ou uma relação de cointegração entre os dados. Opta-se, portanto, pela estimação de um VEC e de um VAR em diferenças com as mesmas variáveis. Por outro lado, os testes realizados para o modelo alternativo 1 mostram sempre uma relação de cointegração, de forma que é estimado apenas um VEC para esse conjunto de dados. Por fim, ressalta-se que o modelo alternativo 2 incorpora uma variável estacionária (serviço da dívida), o que sugere a estimação de um VAR em diferenças. No entanto, de acordo com Kilian e Lütkepohl (2017), também é possível modelar esse conjunto de dados a partir de um VEC com duas relações de cointegração, sendo a primeira relativa ao subsistema de variáveis integradas de ordem 1, e a outra relativa à combinação da primeira relação de cointegração com a variável estacionária. Com efeito, um dos testes de Johansen aponta para a existência de duas relações de cointegração, o que justifica a

estimação do VEC também para essa especificação⁴. As equações obtidas para todos os modelos são mostradas na Tabela 7 do Apêndice A.

Em primeiro lugar, para o modelo VEC base, vale destacar que a equação de longo prazo possui sinais coerentes com o indicado pelos modelos econômicos considerados, com o consumo agregado diretamente proporcional à renda e à riqueza financeira. Ressalta-se que a sensibilidade do consumo a renda de 0,8, na relação de cointegração, é inferior às estimativas mais próximas da unidade obtidas em Lopes (2017). A equação do consumo agregado possui coeficiente de ajustamento para a tendência de longo prazo negativo e significativo, conforme esperado; no entanto, o coeficiente de ajustamento na equação do PIB também é negativo e significativo, o que vai de encontro ao resultado ao resultado de Campbell (1987), que sugere que esse termo deveria ter sinal positivo. Ademais, cabe ressaltar que, nessa especificação, é selecionada apenas uma defasagem para o VAR em nível, de forma que se estima um VEC com zero defasagens. Por fim, o VAR em diferenças, com os mesmos dados, possui sinais esperados na equação do consumo, embora nenhum regressor tenha coeficiente estatisticamente significativo.

Já no modelo VEC alternativo 1, a renda e o crédito possuem os sinais esperados na tendência de longo prazo, mas o *ibovespa* aparece com sinal negativo inesperado. Nessa especificação, o ajuste para a tendência de longo prazo é feito pelo crescimento do PIB e do crédito, o que se aproxima mais do sugerido por Campbell (1987). Ademais, o impacto positivo e estatisticamente significativo da variação do crédito sobre a variação do consumo, no curto prazo, corrobora o arcabouço teórico proposto por Ludvigson (1999), enquanto contrasta com as evidências de menor relevância do crédito apresentadas por Brady (2008) para o caso americano e um período mais recente.

Por fim, a primeira relação de cointegração do VEC alternativo 2 (que inclui o consumo) também possui os sinais esperados, embora com um coeficiente maior para o *ibovespa* do que para o PIB, ao contrário do que ocorre no modelo VEC base. Nesse caso, o ajuste para a tendência de longo prazo é feito, como no modelo base, por meio de uma retração lenta do consumo agregado. Por sua vez, o modelo VAR

⁴ Vale ressaltar que essa especificação apresenta maiores indícios de autocorrelação dos resíduos, mesmo com o aumento do número de defasagens. Contudo, opta-se por seguir com a estimação desse modelo e avaliar a acurácia das projeções geradas fora da amostra.

alternativo 2 sugere que o serviço da dívida possui impacto negativo e estatisticamente significativo sobre o crescimento do consumo em um trimestre à frente. Dessa forma, é possível afirmar que o maior comprometimento da renda do setor privado com dívidas impacta negativamente o consumo das famílias, em linha com os resultados encontrados na literatura e com a intuição econômica.

Estimados os modelos, são realizadas projeções fora da amostra em diferentes horizontes de tempo, com período de avaliação entre 2014 e 2019. Os resultados da razão do RMSE e dos testes de significância estatística em relação aos *benchmarks* (passeio aleatório e AR(1)) são mostrados na Tabela 3.

Tabela 3 – Acurácia das previsões fora da amostra para modelos de curto e médio prazo

Modelo base VEC					
Horizonte	RMSE	RMSE relativo (PA)	RMSE relativo (AR(1))	DM (PA)	DM (AR(1))
1	0,0098	0,92	0,90	4,2%*	9,9%
2	0,0141	0,92	0,85	12,4%	15,0%
4	0,0255	0,96	0,81	29,5%	16,1%
6	0,0279	0,95	0,85	37,7%	31,5%
Modelo base VAR					
Horizonte	RMSE	RMSE relativo (PA)	RMSE relativo (AR(1))	DM (PA)	DM (AR(1))
1	0,0099	0,92	0,90	9,0%	2,7%*
2	0,0157	1,03	0,95	63,2%	18,2%
4	0,0293	1,10	0,92	74,1%	4,9%*
6	0,034	1,16	1,04	75,1%	86,5%
Modelo alternativo 1 VEC (crédito)					
Horizonte	RMSE	RMSE relativo (PA)	RMSE relativo (AR(1))	DM (PA)	DM (AR(1))
1	0,0093	0,87	0,85	1,5%*	2,4%*
2	0,0138	0,91	0,84	6,6%	8,3%
4	0,0256	0,97	0,81	32,6%	12,4%
6	0,0276	0,95	0,85	28,8%	28,3%
Modelo alternativo 2 VEC (serv. dívida)					
Horizonte	RMSE	RMSE relativo (PA)	RMSE relativo (AR(1))	DM (PA)	DM (AR(1))
1	0,0096	0,90	0,88	4,6%*	8,8%
2	0,0137	0,90	0,83	12,4%	14,0%
4	0,0251	0,95	0,79	26,3%	15,4%
6	0,0265	0,91	0,81	29,0%	28,1%
Modelo alternativo 2 VAR (serv. dívida)					
Horizonte	RMSE	RMSE relativo (PA)	RMSE relativo (AR(1))	DM (PA)	DM (AR(1))
1	0,0112	1,05	1,01	67,2%	54,2%
2	0,0174	1,14	1,04	84,6%	67,4%
4	0,0304	1,15	0,94	79,1%	14,3%
6	0,0413	1,42	1,22	87,1%	95,6%

Elaboração própria. *, ** e *** indicam significância a 5%, 1% e 0,1%, respectivamente.

Em geral, as melhores previsões são obtidas a partir dos três modelos VEC, que apresentam RMSE relativo sempre abaixo da unidade, independente do *benchmark* utilizado e do horizonte de previsão. De acordo com o teste de DM, as projeções desses modelos diferem estatisticamente dos *benchmarks* no horizonte de um trimestre à frente, no máximo ao nível de 10%. Em particular, o modelo VEC alternativo 1 também apresenta ganhos de previsão estatisticamente significativos para dois trimestres à frente, em relação a ambos os *benchmarks* e ao nível de 10%, o que corrobora que o crédito às famílias é uma variável útil para projetar o consumo agregado no caso brasileiro. Por sua vez, o modelo VEC alternativo 2, que inclui o serviço da dívida, possui o menor RMSE para três dos quatro horizontes de previsão avaliados.

Por outro lado, as demais especificações apresentam evidências mistas quanto à acurácia preditiva. Em primeiro lugar, o modelo VAR base supera o passeio aleatório apenas para um trimestre à frente, embora o ganho de acurácia preditiva nesse horizonte seja significativo ao nível de 10%. Ademais, o modelo VAR alternativo 2 mostra erro acima dos *benchmarks* para a maior parte dos casos analisados.

Por último, emprega-se o MCS para definir o conjunto que contém os modelos estatisticamente superiores. Os p-valores obtidos, para cada horizonte de previsão, são mostrados na Tabela 8. De forma geral, os dados utilizados se mostram pouco informativos, uma vez que, para os horizontes de projeção de um, dois e quatro trimestres à frente, todas as especificações e os dois *benchmarks* fazem parte do conjunto obtido. No entanto, para seis trimestres à frente, a metodologia exclui o VAR alternativo 2 do conjunto, ao nível de confiança de 85%, reforçando os resultados obtidos na Tabela 3, que mostram elevado erro fora da amostra para essa especificação principalmente no horizonte de seis trimestres à frente.

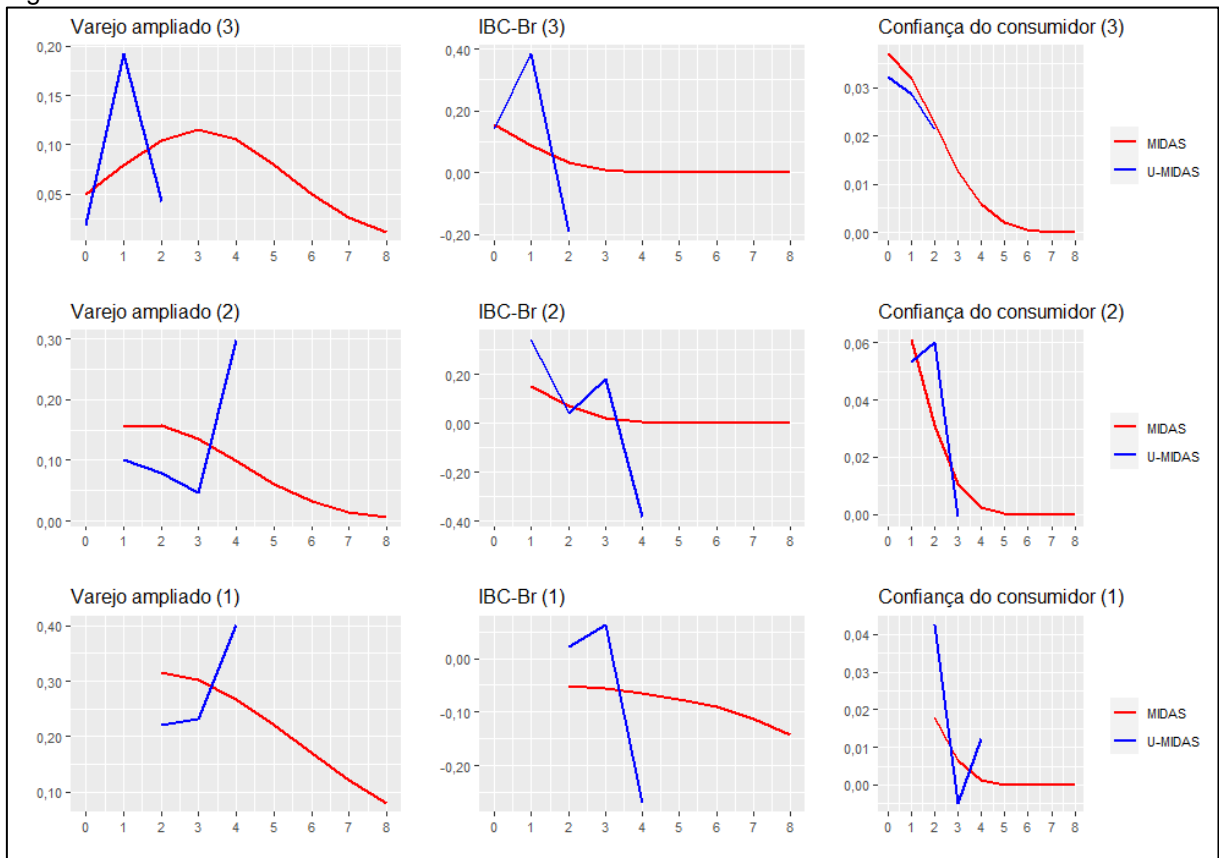
De forma geral, os resultados obtidos apontam que os modelos de correção de erro são úteis para a projeção do consumo agregado, como evidenciado em De Bondt, Gieseck e Zekaite (2020). Ademais, o bom desempenho preditivo do VEC alternativo 1 se aproxima do resultado obtido por Schettini et al. (2012), segundo o qual o modelo de cointegração, contendo apenas a renda disponível e o crédito às famílias, permite a maior redução do erro quadrático médio fora da amostra, na comparação às demais especificações utilizadas.

4.2 CURTÍSSIMO PRAZO

Assim como na seção anterior, o primeiro passo da análise consiste na identificação da ordem de integração de todas as variáveis utilizadas no modelo de curtíssimo prazo (consumo agregado, vendas no varejo ampliado, IBC-Br e confiança do consumidor). A Tabela 9, no Apêndice B, mostra o resultado do teste ADF. Todas as variáveis são integradas de ordem um, de forma que são incorporadas ao modelo em primeira diferença.

Na sequência, busca-se definir a melhor especificação para o MIDAS restrito, assim como o melhor irrestrito (U-MIDAS), tomando novamente a amostra até o quarto trimestre de 2013. Para cada caso, são estimados três modelos, incorporando gradativamente os dados mensais contemporâneos ao trimestre previsto. A Figura 3 apresenta os coeficientes estimados, para cada variável, utilizando os modelos MIDAS e U-MIDAS. O número de meses tomados como conhecidos por cada especificação aparece entre parênteses. As Tabelas 10, 11 e 12, no Apêndice B, mostram os coeficientes estimados para cada modelo e variável, além dos parâmetros associados aos polinômios dos modelos MIDAS. Cabe ressaltar que o teste de hipóteses é realizado somente para os coeficientes dos modelos U-MIDAS e para os parâmetros associados aos polinômios dos modelos MIDAS.

Figura 3 – Coeficientes dos modelos MIDAS e U-MIDAS



Elaboração própria.

Para os modelos MIDAS, na maior parte dos casos os coeficientes estimados apresentam sinal positivo e tendência de decaimento, de forma que as defasagens mais altas possuem impacto relativamente mais baixo. A primeira exceção ocorre para os coeficientes do IBC-Br, na especificação com apenas um mês conhecido dentro do trimestre previsto, que apresentam sinais negativos e impacto crescente, em módulo, em maiores defasagens; a segunda exceção ocorre para os coeficientes do varejo ampliado, com todos os dados mensais conhecidos, que apresentam estimativas mais elevadas em defasagens intermediárias. Já nos modelos U-MIDAS, a maior parte dos coeficientes também apresenta o esperado sinal positivo, embora poucas estimativas sejam estatisticamente significativas, o que provavelmente decorre do tamanho da amostra limitado nessa primeira estimação. Vale destacar, no entanto, que as vendas no varejo ampliado parecem ter maior capacidade de explicar o consumo agregado, em relação aos demais regressores utilizados.

Por fim, a adequação dos modelos estimados é avaliada com base nas projeções fora da amostra, novamente na comparação ao passeio aleatório e ao AR(1). Uma vez que a Figura 3 mostra coeficientes distintos para o MIDAS e o U-

MIDAS, mesmo para defasagens iguais, é razoável esperar que as projeções obtidas também serão divergentes, de forma que a avaliação da previsão fora da amostra permitirá discernir qual metodologia possui desempenho superior para o caso analisado. A Tabela 4 mostra, para cada modelo, o RMSE, a razão dessa métrica em relação aos *benchmarks*, e o p-valor do teste de DM.

Tabela 4 – Acurácia das previsões fora da amostra para modelos de curtíssimo prazo

	MIDAS (3)	MIDAS (2)	MIDAS (1)
RMSE	0,0079	0,0062	0,0079
RMSE relativo (PA)	0,75	0,58	0,75
RMSE relativo (AR(1))	0,67	0,52	0,68
DM (PA)	1,1%*	0,1%***	1,3%*
DM (AR(1))	4,5%*	1%*	3%*
	U-MIDAS (3)	U-MIDAS (2)	U-MIDAS (1)
RMSE	0,0109	0,0076	0,0086
RMSE relativo (PA)	1,03	0,71	0,81
RMSE relativo (AR(1))	0,93	0,64	0,72
DM (PA)	59,5%	0,6%**	4,5%*
DM (AR(1))	30,2%	2,3%*	4%*

Elaboração própria. *, ** e *** indicam significância a 5%, 1% e 0,1%, respectivamente.

Em primeiro lugar, vale notar que quase todas as especificações, com exceção do U-MIDAS(3), apresentam razão do RMSE consideravelmente abaixo da unidade na comparação aos dois *benchmarks*, o que sugere boa acurácia preditiva dos modelos propostos. Ademais, vale destacar que, em todos os casos, o modelo MIDAS possui menor RMSE em relação ao U-MIDAS estimado com informação equivalente, em linha com a conclusão de Ferreira (2020). Por fim, é possível afirmar que, à exceção do U-MIDAS(3), todos os modelos estimados possuem erro de previsão estatisticamente inferior a ambos os *benchmarks* empregados, ao nível de 5,0%.

Um resultado inesperado na Tabela 4 é o menor RMSE fora da amostra quando utilizados regressores conhecidos dois meses dentro do trimestre previsto, em comparação à especificação que utiliza a totalidade dos dados contemporâneos. No entanto, algumas evidências análogas são obtidas por Dreger e Kholodilin (2013), que mostram que, para regressores específicos, a melhor previsão é feita com dados conhecidos apenas no primeiro mês do trimestre previsto.

O último passo desse trabalho consiste no uso do MCS para determinar o conjunto que contém os modelos estatisticamente superiores a partir de um grupo

inicial formado por todas as especificações MIDAS e U-MIDAS estimadas, além dos dois *benchmarks* empregados. Os p-valores obtidos são mostrados na Tabela 13. Em particular, essa metodologia aponta que os modelos U-MIDAS(3), passeio aleatório e AR(1) não pertencem ao conjunto final obtido, em linha com as leituras do erro fora da amostra, ao nível de confiança de 85%.

5 CONCLUSÃO

Esse trabalho avaliou a previsibilidade da série brasileira do consumo agregado em diferentes horizontes de tempo. Com foco em projeções de curto e médio prazo, foram empregados os modelos VEC e VAR, com determinantes de renda, riqueza financeira, crédito e serviço da dívida. Já para o horizonte de curtíssimo prazo (*nowcasting*), foi utilizada a metodologia MIDAS, tendo como regressores as séries de vendas no varejo ampliado, confiança do consumidor e IBC-Br.

Na primeira parte do trabalho, com foco no curto e médio prazo, os modelos VEC estimados apresentam leve redução do RMSE em todos os horizontes de projeção fora da amostra, na comparação ao modelo de passeio aleatório e ao AR(1), embora só sejam observados ganhos estatisticamente significantes de acurácia preditiva para um trimestre à frente (ao nível de 5%). Ademais, é possível destacar que os modelos de correção de erros que incorporam o crédito às famílias ou o serviço da dívida mostram ganhos de acurácia preditiva em relação às demais especificações estimadas. Em síntese, a relevância dessas variáveis aponta, ao mesmo tempo, para maior sensibilidade do consumo agregado a choques macroeconômicos e para maior eficácia da atuação da política econômica sobre a atividade no curto prazo, em especial para medidas que influenciem o balanço das famílias através do crédito.

Já na segunda parte do trabalho, de forma geral, a especificação MIDAS mostra expressiva melhora da acurácia preditiva em comparação aos *benchmarks*. Em primeiro lugar, é possível concluir que o conjunto de dados de alta frequência empregado é relevante para projetar o consumo agregado, com destaque para as vendas no varejo ampliado, reforçando a relevância dessa metodologia para a avaliação tempestiva da conjuntura macroeconômica. Além disso, os resultados obtidos apontam para a superioridade do MIDAS em relação ao U-MIDAS para cada conjunto de informação, o que evidencia os ganhos de empregar uma restrição polinomial que combine flexibilidade e parcimônia na estimação dos coeficientes.

Uma possível extensão desse trabalho consiste na incorporação de novos regressores correlacionados à variação do consumo agregado. Para o horizonte de curto e médio prazo, um potencial preditor útil seria o preço dos imóveis residenciais

(riqueza imobiliária); já no curtíssimo prazo, uma variável importante seria um índice dos serviços prestados às famílias. Vale destacar, no entanto, que a indisponibilidade de séries temporais longas o suficiente para representar essas variáveis consiste em um entrave importante para a modelagem do consumo agregado das famílias no Brasil.

REFERÊNCIAS

- ALBUQUERQUE, Bruno; KRUSTEV, Georgi. *Debt overhang and deleveraging in the US household sector: gauging the impact on consumption*. **Review of Income and Wealth**, v. 64, n. 2, p. 459-481, 2018.
- AL-EYD, Ali et al. *Part 1: assessing the factors of resilience of private consumption in the euro area—a macroeconomic perspective*. In: DERROOSE, Servaas. **Assessing the factors of resilience of private consumption in the euro area**, European Commission Directorate, General for Economic and Financial Affairs Economic papers, n. 252, p. 1–94, 2006.
- ARMESTO, Michelle T.; ENGEMANN, Kristie M; OWYANG, Michael T. *Forecasting with mixed frequencies*. **Federal Reserve Bank of St. Louis Review**, v. 92, n. 6, p. 521-36, 2010.
- BANCO CENTRAL DO BRASIL. **Relatório de Inflação**, v. 3, n. 3, 2001. Disponível em: <https://www.bcb.gov.br/content/ri/relatorioinflacao/200109/RELINF200109-ri200109P.pdf>. Acesso em: 25 out. 2022.
- BRADY, Ryan R. *Structural breaks and consumer credit: Is consumption smoothing finally a reality?* **Journal of macroeconomics**, v. 30, n. 3, p. 1246-1268, 2008.
- BERNARDI, Mauro; CATANIA, Leopoldo. *The model confidence set package for R*. **International Journal of Computational Economics and Econometrics**, v. 8, n. 2, p. 144-158, 2018.
- CAMPBELL, John Y. *Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis*. **Econometrica**, v. 55, p. 1249-74, 1987.
- DE BONDT, Gabe J.; GIESECK, Arne; ZEKAITE, Zivile. *Thick modelling income and wealth effects: a forecast application to euro area private consumption*. **Empirical Economics**, v. 58, n. 1, p. 257-286, 2020.
- DICKEY, David A.; FULLER, Wayne A. *Distribution of the estimators for autoregressive time series with a unit root*. **Journal of the American statistical association**, v. 74, n. 366a, p. 427-431, 1979.
- DIEBOLD, Francis X.; MARIANO, Roberto S. *Comparing Predictive Accuracy*. **Journal of Business and Economic Statistics**, v. 13, p. 253-265, 1995.
- DREGER, Christian; KHOLODILIN, Konstantin A. *Forecasting private consumption by consumer surveys*. **Journal of Forecasting**, v. 32, n. 1, p. 10-18, 2013.
- DUARTE, Cláudia; RODRIGUES, Paulo MM; RUA, António. *A mixed frequency approach to the forecasting of private consumption with ATM/POS data*. **International Journal of Forecasting**, v. 33, n. 1, p. 61-75, 2017.
- ENGLE, Robert F.; GRANGER, Clive WJ. *Co-integration and error correction: representation, estimation, and testing*. **Econometrica: journal of the Econometric Society**, p. 251-276, 1987.

ESTRADA, Ángel et al. *Household debt and uncertainty: Private consumption after the Great Recession*. **Banco de España**, n. 1415, 2014.

FELINI, Patrícia; GOMES, Fábio A. R.; SOAVE, Gian P. Previsão do Consumo Agregado: o papel de índices de confiança do consumidor. **Estudos Econômicos**, v. 52, p. 243-279, 2022.

FERNANDEZ-CORUGEDO, Emilio; PRICE, Simon; BLAKE, Andrew. *The dynamics of consumers' expenditure: the UK consumption ECM redux*. **Bank of England**, 2003.

FERREIRA, Julia L. ***Transmuting unequally spaced data: a MIDAS regression touch to forecast real GDP growth in Brazil***. 2020. Dissertação (Mestrado em Economia) – Escola de Economia de São Paulo, Fundação Getulio Vargas, São Paulo, 2020.

FORONI, Claudia; MARCELLINO, Massimiliano. *A comparison of mixed frequency approaches for nowcasting Euro area macroeconomic aggregates*. **International Journal of Forecasting**, v. 30, n. 3, p. 554-568, 2014.

GHYSELS, Eric; SANTA-CLARA, Pedro; VALKANOV, Rossen. *The MIDAS touch: Mixed data sampling regression models*. **Cirano Working Papers**, 2004.

GHYSELS, Eric; SINKO, Arthur; VALKANOV, Rossen. *MIDAS regressions: Further results and new directions*. **Econometric reviews**, v. 26, n. 1, p. 53-90, 2007.

GHYSELS, Eric; KVEDARAS, Virmantas; ZEMLYS, Vaidotas. *Mixed frequency data sampling regression models: the R package midasr*. **Journal of statistical software**, v. 72, p. 1-35, 2016.

GOMES, Fábio A. R. Consumo no Brasil: comportamento otimizador, restrição de crédito ou miopia? **Revista brasileira de Economia**, v. 64, p. 261-275, 2010.

HALL, Robert. *Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence*. **Journal of Political Economy**, v. 86, n. 6, p. 971-987, 1978.

HANSEN, Peter R.; LUNDE, Asger; NASON, James M. *The model confidence set*. **Econometrica**, v. 79, n. 2, p. 453-497, 2011.

HARVEY, David; LEYBOURNE, Stephen; NEWBOLD, Paul. *Testing the equality of prediction mean squared errors*. **International Journal of forecasting**, v. 13, n. 2, p. 281-291, 1997.

IBGE. **Contas nacionais trimestrais**, 2022. Disponível em: https://agenciadenoticias.ibge.gov.br/media/com_mediaibge/arquivos/5bebef228dda5994c8deeb05945e2963.pdf. Acesso em: 01 set. 2022.

JOHANSEN, Søren. *Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models*. **Econometrica: journal of the Econometric Society**, p. 1551-1580, 1991.

JOHANSEN, Søren. *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford: Oxford University Press, 1995.

KILIAN, Lutz; LÜTKEPOHL, Helmut. *Structural vector autoregressive analysis*. Cambridge: Cambridge University Press, 2017.

KIM, Yun K.; SETTERFIELD, Mark; MEI, Yuan. *Aggregate consumption and debt accumulation: an empirical examination of US household behaviour*. **Cambridge Journal of Economics**, v. 39, n. 1, p. 93-112, 2015.

LETTAU, Martin; LUDVIGSON, Sydney. *Consumption, aggregate wealth, and expected stock returns*. **The Journal of Finance**, v. 56, n. 3, p. 815-849, 2001.

LOPES, Luckas S. Testando teorias para o consumo agregado no Brasil. **Nova economia**, v. 27, p. 209-240, 2017.

LUDVIGSON, Sydney; STEINDEL, Charles. *How important is the stock market effect on consumption?* **Economic Policy Review**, v. 5, p. 29-51, 1998.

LUDVIGSON, Sydney. *Consumption and credit: a model of time-varying liquidity constraints*. **Review of economics and statistics**, v. 81, n. 3, p. 434-447, 1999.

PFÄFF, Bernhard. *VAR, SVAR and SVEC models: Implementation within R package vars*. **Journal of statistical software**, v. 27, p. 1-32, 2008.

SAX, Christoph; EDELBUETTEL, Dirk. *Seasonal adjustment by x-13arima-seats in R*. **Journal of Statistical Software**, v. 87, p. 1-17, 2018.

SCHETTINI, Bernardo P. et al. Novas evidências empíricas sobre a dinâmica trimestral do consumo agregado das famílias brasileiras no período 1995-2009. **Economia e Sociedade**, v. 21, p. 607-641, 2012.

STIGLER, Matthieu. *Nonlinear time series in R: Threshold cointegration with tsDyn*. In: RAO, C. Radhakrishna; VINOD, Hrishikesh. **Handbook of Statistics**. Amsterdam: Elsevier, p. 229-264, 2020.

ZUANAZZI, Pedro T.; ZIEGELMANN, Flávio A. Previsões para o crescimento do PIB trimestral brasileiro com séries financeiras e econômicas mensais: uma aplicação de MIDAS. **Economia Aplicada**, v. 18, p. 295-318, 2014.

APÊNDICES

APÊNDICE A – Testes estatísticos e resultados dos modelos de curto e médio prazo

Tabela 5 – Teste ADF para variáveis dos modelos de curto e médio prazo

Variável em nível	Consumo agregado	PIB	Ibovespa	Saldo de crédito	Serviço da dívida
Constante/ tendência	-	-	-	-	tendência
Estatística	-1,40	-2,02	-2,09	2,44	-3,68
Valor crítico (5%)	-3,45	-3,45	-2,89	-1,95	-3,45
Primeira diferença	Consumo agregado	PIB	Ibovespa	Saldo de crédito	Serviço da dívida
Constante/ tendência	ambos	constante	-	constante	-
Estatística	-5,86	-6,28	-5,90	-4,48	-4,14
Valor crítico (5%)	-3,45	-2,89	-1,95	-2,89	-1,95
Ordem de integração	I (1)	I (1)	I (1)	I (1)	I (0)

Elaboração própria.

Tabela 6 – Procedimento de Johansen para modelos de curto e médio prazo

Traço	Modelo base		Alternativo 1 (crédito)		Alternativo 2 (serv. dívida)	
	Sem constante	Com constante	Sem constante	Com constante	Sem constante	Com constante
Nenhuma	1,3%*	9,88%	0%***	0,01%***	0%***	0,1%**
No máximo uma	31,52%	32,89%	19,65%	36,94%	6,94%	35,50%
No máximo duas	31,56%	46,12%	78,30%	93,36%	63,30%	39,41%
No máximo três	-	-	93,09%	58,86%	69,33%	29,45%
Máximo autovalor	Modelo base		Alternativo 1 (crédito)		Alternativo 2 (serv. dívida)	
	Sem constante	Com constante	Sem constante	Com constante	Sem constante	Com constante
Nenhuma	1,16%*	14,12%	0%***	0%***	0%***	0,04%***
No máximo uma	44,28%	29,75%	8,51%	16,18%	3,09%*	51,40%
No máximo duas	31,56%	46,12%	62,83%	92,47%	59,02%	42,09%
No máximo três	-	-	93,09%	58,86%	69,33%	29,45%

Elaboração própria. *, ** e *** indicam significância a 5%, 1% e 0,1%, respectivamente.

Tabela 7 – Estimativas para modelos de curto e médio prazo

Modelo base VEC				
Relação de cointegração	Consumo (t) = 0,22 + 0,80 * PIB (t) + 0,15 * Ibovespa (t)			
	Δ Consumo (t)	Δ PIB (t)	Δ Ibovespa (t)	
Correção de erro (t-1)	-0,0680***	-0,0633**	0,1289	
Modelo base VAR				
	Δ Consumo (t)	Δ PIB (t)	Δ Ibovespa (t)	
Constante	0,0038*	0,0044*	0,0231	
Δ Consumo (t-1)	0,0937	0,3000	1,4882	
Δ PIB (t-1)	0,0973	-0,1142	-3,6133**	
Δ Ibovespa (t-1)	0,0176	0,0341*	0,3232**	
Modelo alternativo 1 VEC (crédito)				
Relação de cointegração	Consumo (t) = 2,87 + 0,39 * PIB (t) - 0,10 * Ibovespa (t) + 0,24 * Crédito (t)			
	Δ Consumo (t)	Δ PIB (t)	Δ Ibovespa (t)	Δ Crédito (t)
Correção de erro (t-1)	-0,0325	0,1690*	-0,5725	0,7354***
Δ Consumo (t-1)	0,1429	0,1985	2,1734	0,0595
Δ PIB (t-1)	0,0321	-0,153	-3,5067**	0,1150
Δ Ibovespa (t-1)	0,0171	0,0257	0,3379**	-0,0120
Δ Crédito (t-1)	0,1460***	0,0618	0,5370	0,1956***
Modelo alternativo 2 VEC (serv. dívida)				
Relação de cointegração 1	Consumo (t) = 1,96 + 0,26 * PIB (t) + 0,55 * Ibovespa (t)			
Relação de cointegração 2	Serviço da dívida (t) = 0,76 + 1,30 * PIB (t) - 1,35 * Ibovespa (t)			
	Δ Consumo (t)	Δ PIB (t)	Δ Ibovespa (t)	Δ Serviço da dívida (t)
Correção de erro 1 (t-1)	-0,0328*	-0,0188	0,2698	-0,5167***
Correção de erro 2 (t-1)	-0,0031	0,0053	0,1549	-0,3232***
Modelo alternativo 2 VAR (serv. dívida)				
	Δ Consumo (t)	Δ PIB (t)	Δ Ibovespa (t)	Δ Serviço da dívida (t)
Constante	0,0646**	0,0393	-0,1505	0,3330*
Δ Consumo (t-1)	-0,1224	0,2201	0,1107	0,4364
Δ PIB (t-1)	-0,0588	-0,1650	-2,6709	0,8022
Δ Ibovespa (t-1)	0,0321*	0,0319	0,3239*	-0,2894**
Δ Serviço da dívida (t-1)	-0,0199**	-0,0114	0,0609	0,8776***

Elaboração própria. *, ** e *** indicam significância a 5%, 1% e 0,1%, respectivamente.

Tabela 8 – P-valores do MCS para cada horizonte de previsão do curto e médio prazo

	1	2	4	6
Modelo base VEC	100,0%	100,0%	100,0%	100,0%
Modelo base VAR	100,0%	85,8%	64,7%	38,3%
Modelo alternativo 1 VEC (crédito)	100,0%	100,0%	100,0%	100,0%
Modelo alternativo 2 VEC (serv. dívida)	100,0%	100,0%	100,0%	100,0%
Modelo alternativo 2 VAR (serv. dívida)	62,7%	33,7%	35,8%	-
AR(1)	42,2%	49,8%	29,3%	36,8%
Passeio aleatório	39,6%	100,0%	100,0%	100,0%

Elaboração própria.

APÊNDICE B – Testes estatísticos e resultados dos modelos de curtíssimo prazo

Tabela 9 – Teste ADF para variáveis dos modelos de curtíssimo prazo

Variável em nível	Consumo agregado	Varejo ampliado	IBC-Br	Confiança do consumidor
Constante/ tendência	-	-	constante	-
Estatística	3,80	5,57	-2,28	-3,36
Valor crítico (5%)	-1,95	-1,95	-3,43	-3,43
Primeira diferença	Consumo agregado	Varejo ampliado	IBC-Br	Confiança do consumidor
Constante/ tendência	constante	constante	constante	-
Estatística	-5,49	-11,10	-7,97	-11,35
Valor crítico (5%)	-2,88	-2,88	-2,88	-1,95
Ordem de integração	I (1)	I (1)	I (1)	I (1)

Elaboração própria.

Tabela 10 – Coeficientes dos modelos MIDAS

	MIDAS (3)	MIDAS (2)	MIDAS (1)
	Δ Consumo (t)	Δ Consumo (t)	Δ Consumo (t)
Δ Varejo ampliado(t)	0,0496	-	-
Δ Varejo ampliado (t-1/3)	0,0792	0,1560	-
Δ Varejo ampliado (t-2/3)	0,1049	0,1570	0,3157
Δ Varejo ampliado(t-1)	0,1153	0,1345	0,3016
Δ Varejo ampliado(t-4/3)	0,1052	0,0980	0,2678
Δ Varejo ampliado(t-5/3)	0,0796	0,0607	0,2210
Δ Varejo ampliado(t-2)	0,0500	0,0320	0,1695
Δ Varejo ampliado(t-7/3)	0,0261	0,0144	0,1208
Δ Varejo ampliado(t-8/3)	0,0113	0,0055	0,0800
Δ Confiança do cons. (t)	0,0369	-	-
Δ Confiança do cons. (t-1/3)	0,0321	0,0611	-
Δ Confiança do cons. (t-2/3)	0,0225	0,0312	0,0179
Δ Confiança do cons. (t-1)	0,0128	0,0105	0,0066
Δ Confiança do cons. (t-4/3)	0,0058	0,0023	0,0013
Δ Confiança do cons. (t-5/3)	0,0022	0,0003	0,0001
Δ Confiança do cons. (t-2)	0,0006	0,0000	0,0000
Δ Confiança do cons. (t-7/3)	0,0002	0,0000	0,0000
Δ Confiança do cons. (t-8/3)	0,0000	0,0000	0,0000
Δ IBC-Br (t)	0,1538	-	-
Δ IBC-Br (t-1/3)	0,0888	0,1494	-
Δ IBC-Br (t-2/3)	0,0332	0,0695	-0,0510
Δ IBC-Br (t-1)	0,0080	0,0204	-0,0565
Δ IBC-Br (t-4/3)	0,0012	0,0038	-0,0644
Δ IBC-Br (t-5/3)	0,0001	0,0004	-0,0755
Δ IBC-Br (t-2)	0,0000	0,0000	-0,0910
Δ IBC-Br (t-7/3)	0,0000	0,0000	-0,1127
Δ IBC-Br (t-8/3)	0,0000	0,0000	-0,1435

Elaboração própria.

Tabela 11 – Coeficientes estimados para os modelos U-MIDAS

	U-MIDAS (3)	U-MIDAS (2)	U-MIDAS (1)
	Δ Consumo (t)	Δ Consumo (t)	Δ Consumo (t)
Intercepto	0,0086***	0,0082**	0,0072**
Δ Varejo ampliado(t)	0,0193	-	-
Δ Varejo ampliado (t-1/3)	0,1919	0,1017	-
Δ Varejo ampliado (t-2/3)	0,0422	0,0794	0,2203
Δ Varejo ampliado(t-1)	-	0,0466	0,2318
Δ Varejo ampliado(t-4/3)	-	0,2992*	0,4016*
Δ Confiança do cons. (t)	0,0322	-	-
Δ Confiança do cons. (t-1/3)	0,0286	0,0532	-
Δ Confiança do cons. (t-2/3)	0,0212	0,0601	0,0425
Δ Confiança do cons. (t-1)	-	-0,0011	-0,0048
Δ Confiança do cons. (t-4/3)	-	-	0,0122
Δ IBC-Br (t)	0,1438	-	-
Δ IBC-Br (t-1/3)	0,3834	0,3419	-
Δ IBC-Br (t-2/3)	-0,1927	0,0423	0,0212
Δ IBC-Br (t-1)	-	0,1804	0,0627
Δ IBC-Br (t-4/3)	-	-0,3864	-0,2698

Elaboração própria. *, ** e *** indicam significância a 5%, 1% e 0,1%, respectivamente.

Tabela 12 – Parâmetros estimados para os modelos MIDAS

	MIDAS (3)	MIDAS (2)	MIDAS (1)
β_0	0,0071**	0,0067**	0,0050
Varejo ampliado (β_1)	0,6210	0,6581	1,4763**
Varejo ampliado (θ_1)	0,7473	0,2494	0,0642
Varejo ampliado (θ_2)	-0,0933	-0,0808	-0,0366*
Confiança do consumidor (β_1)	0,1131	0,1055	0,0258
Confiança do consumidor (θ_1)	0,1791	-0,0438	-0,0337
Confiança do consumidor (θ_2)	-0,1067	-0,2097	-0,3232
IBC-Br (β_1)	0,2852	0,2436	-0,5945
IBC-Br (θ_1)	0,1061	-0,0773	0,0617
IBC-Br (θ_2)	-0,2183	-0,2294	0,0139

Elaboração própria. *, ** e *** indicam significância a 5%, 1% e 0,1%, respectivamente.

Tabela 13 – P-valores do MCS para o curtíssimo prazo

	P-valor
MIDAS (3)	99,9%
U-MIDAS (3)	-
MIDAS (2)	100,0%
U-MIDAS (2)	100,0%
MIDAS (1)	87,3%
U-MIDAS (1)	17,8%
AR(1)	-
Passeio aleatório	-

Elaboração própria.