



Thiago Guedes Morais

**Fluxos de capitais para economias
emergentes: o caso do Brasil**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Macroeconomia e Finanças do Departamento de Economia do Centro de Ciências Sociais da PUC-Rio.

Orientador: Prof. Márcio Gomes Pinto Garcia

Rio de Janeiro

Março de 2022



Thiago Guedes Morais

**Fluxos de capitais para economias
emergentes: o caso do Brasil**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Macroeconomia e Finanças do Departamento de Economia do Centro de Ciências Sociais da PUC-Rio. Aprovada pela Comissão Examinadora abaixo:

Prof. Márcio Gomes Pinto Garcia

Orientador

Departamento de Economia – PUC-Rio

Prof. Marcelo Cunha Medeiros

Departamento de Economia – PUC-Rio

Prof. Waldyr Dutra Aersa

Departamento de Economia – PUC-Rio

Rio de Janeiro, 28 de Março de 2022

Todos os direitos reservados. A reprodução, total ou parcial, do trabalho é proibida sem autorização da universidade, da autora e do orientador.

Thiago Guedes Morais

Graduou-se em engenharia mecânica pela Universidade Federal do Rio de Janeiro (UFRJ) em 2018.

Ficha Catalográfica

Morais, Thiago Guedes

Fluxos de capitais para economias emergentes: o caso do Brasil / Thiago Guedes Morais; orientador: Márcio Gomes Pinto Garcia. – 2022.

33 f.: il. color. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia.

Inclui bibliografia

1. Economia – Teses. 2. Fluxos de capitais. 3. Balanço de Pagamentos. 4. Conta Financeira. 5. *Machine learning*. 6. Regularização. I. Garcia, Márcio Gomes Pinto. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. III. Título.

CDD: 330

Agradecimentos

Aos meus pais, Claudia e Marcos, ao meu irmão, Matheus, e à minha família, pelo amor, incentivo e apoio incondicional em todos os momentos.

À minha esposa, Renata, pelo carinho, compreensão e suporte.

Ao meu avô, Pecu, pela inspiração e apoio em buscar meus sonhos.

Aos meus amigos, pela compreensão nos momentos de ausência e, em especial ao Fernando Secco, pelas palavras de calma e incentivo, essenciais para conclusão deste trabalho.

Aos meus colegas de trabalho, em especial ao Henrique Pires, pela paciência, explicações e ajuda no desenvolvimento dessa dissertação.

Ao professor, Márcio Garcia, pela disponibilidade e orientação na conclusão do mestrado.

Resumo

Morais, Thiago Guedes; Garcia, Márcio Gomes Pinto. **Fluxos de capitais para economias emergentes: o caso do Brasil**. Rio de Janeiro, 2022. 33p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Motivados pela posição de destaque do real brasileiro entre uma das moedas mais depreciadas em relação aos seus pares emergentes em meados de 2020, potencialmente fomentada pela expressiva evasão de capitais observada no decorrer da pandemia COVID-19 que culminou com um déficit no mercado cambial, realizamos previsões um trimestre a frente para os fluxos de capitais líquidos para o Brasil através de técnicas de *machine learning*, utilizando modelos de regularização para seleção das variáveis importantes. Os fluxos são obtidos a partir de dados trimestrais do balanço de pagamentos, englobando 2004:T1 a 2021:T1. Os modelos propostos, tanto LASSO quanto *adaLASSO* + OLS, foram capazes de gerar previsões fora da amostra melhores que o modelo de *benchmark*, AR. Apesar disso, quando comparados entre si, não podemos rejeitar a hipótese nula de que os modelos propostos possuem a mesma precisão de previsão.

Palavras-chave

Fluxos de capitais; Balanço de Pagamentos; Conta Financeira; *Machine learning*; Regularização

Abstract

Morais, Thiago Guedes; Garcia, Márcio Gomes Pinto. **Capital flows to emerging markets: the case of Brazil**. Rio de Janeiro, 2022. 33p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Motivated by the prominent position of the Brazilian real among the most depreciated currencies in comparison with its emerging peers in mid-2020, potentially fueled by the significant capital outflow observed during the COVID-19 pandemic that resulted in a deficit in the foreign exchange market, we make one quarter-ahead forecast for net capital flows to Brazil through machine learning techniques, using shrinkage methods to select important variables. These flows are computed from quarterly balance of payments data from 2004:Q1 to 2021:Q1. The proposed models, both LASSO and adaLASSO + OLS, were able to generate better out-of-sample forecasts than the benchmark model, AR. Nevertheless, when compared to each other, we cannot reject the null hypothesis that the proposed models have the same forecast accuracy.

Keywords

Capital flows; Balance of Payments; Financial Account; Machine learning; Shrinkage

Sumário

Introdução	10
Revisão da literatura	14
Dados	17
Estratégia empírica	20
Modelo empírico	21
Resultados	24
Conclusão	29
Referências	30
Anexo 1 – Conjuntos estimados	32

Lista de figuras

Figura 1 - Taxa de câmbio efetiva real. Fonte: Bank for International Settlements	11
Figura 2 - Transações Correntes vs. Conta Financeira. Fonte: Banco Central do Brasil	11
Figura 3 - Conta Financeira. Fonte: Banco Central do Brasil	12
Figura 4 - Conta Financeira em US\$ bilhões vs. CDS (eixo direito, invertido)	20
Figura 5 - Janela crescente ou rolling forward	21
Figura 6 - Dispersão Conta Financeira vs. CDS	24
Figura 7 - Erros de previsão dos modelos fora da amostra em US\$ bilhões	27

Lista de tabelas

Tabela 1 - Dados	18
Tabela 2 - Estatísticas descritivas	19
Tabela 3 - Média da raiz do erro quadrático médio e do erro médio absoluto	26
Tabela 4 - P-valor dos testes Diebold-Mariano	28
Tabela 5 - Raiz do erro quadrático médio e erro médio absoluto do conjunto selecionado	28

Introdução

De acordo com a teoria econômica, a livre circulação de capitais possibilita a alocação eficiente dos recursos, aumentando produtividade e crescimento econômico (Ahmed e Zlate, 2014). Historicamente, ondas de influxo de capitais em países emergentes foram acompanhadas de acumulação de reservas, apreciação do câmbio real, expansão no mercado de ações, aumento do crescimento econômico e expansão do déficit de transações correntes (Calvo, Leiderman e Reinhart, 1993). Por outro lado, esses países também sofreram uma forte evasão de capitais em momentos de maior incerteza, sendo vítimas recorrentes da “fuga para a qualidade” quando há maior aversão ao risco e preferência por liquidez.

No Brasil, foco do presente trabalho, as transações econômicas entre o país e o exterior (entre residentes e não-residentes) - que compreendem a balança comercial de bens e serviços, as contas de renda primária e secundária, a conta de capital e a conta financeira - são registradas no Balanço de Pagamentos (BP), que segue as diretrizes conceituais do Fundo Monetário Internacional (FMI) estabelecidas no documento “*Balance of Payments and International Investment Position Manual*” (BPM6, 2011). Sua estrutura resume-se a três grandes categorias: as transações correntes (TC), a conta de capital (CC) e a conta financeira (CF). Sendo o objeto de interesse do presente trabalho, cabe mencionar que a conta financeira é dividida em quatro grupos: investimento direto, investimentos em carteira, derivativos e outros investimentos. Cada grupo desdobra-se em ativos e passivos, refletindo os fluxos envolvendo ativos externos detidos por residentes e ativos domésticos detidos por não-residentes, respectivamente. Por fim, o somatório de transações correntes e da conta de capital representa o empréstimo líquido (superávit) ou endividamento líquido (déficit) do país em relação ao resto do mundo, que é conceitualmente igual ao resultado líquido da conta financeira mais a variação nos ativos de reserva. Portanto, em outras palavras, as duas últimas representam a forma de financiamento das duas primeiras.

De forma geral, é comum que as atenções se voltem para o resultado em transações correntes nas divulgações do BP pelo Banco Central. Por muitas vezes, a análise das fontes de financiamento é preterida. Sob essa perspectiva, com a posição de destaque entre uma das moedas mais depreciadas em relação aos seus pares emergentes em meados de 2020 com o decorrer da crise sanitária COVID-19 (Figura 1), não era incomum a percepção de um certo otimismo por parte de operadores de câmbio a respeito do real brasileiro. Contribuía para isso o cenário de atividade fraca (via redução de importações), câmbio depreciado (via aumento de competitividade), preços de commodities em alta (via melhora no saldo de balança comercial) e restrições de deslocamento impostas pela pandemia (via redução de despesas com viagens internacionais), que ajustava com rigidez o déficit em transações correntes, reduzindo a necessidade de financiamento do país (Figura 2). Apesar desse ajuste contribuir para um balanço cambial superavitário, ele não é suficiente para garanti-lo, isto é:

$$\text{Balanço cambial (Ativos de Reserva)} = TC + CC - CF$$

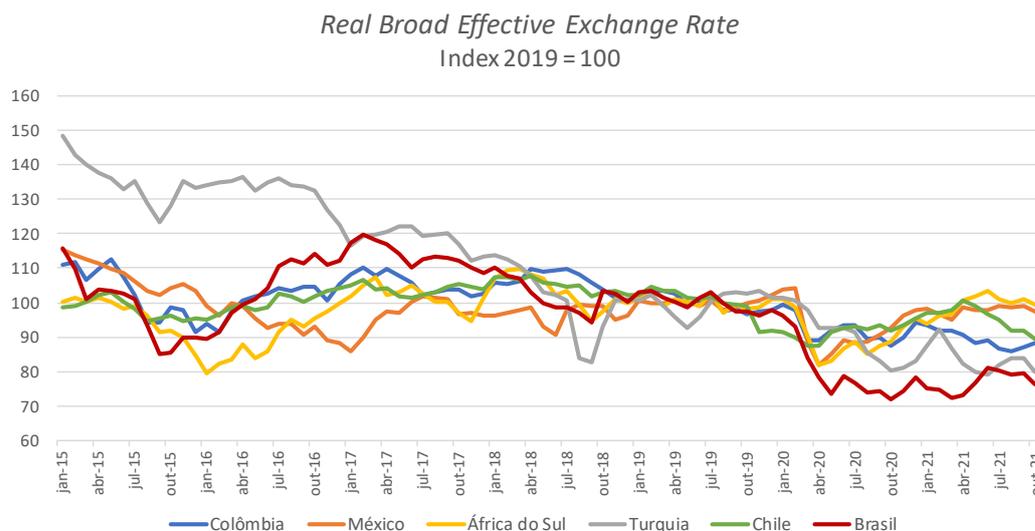


Figura 1 - Taxa de câmbio efetiva real. Fonte: Bank for International Settlements

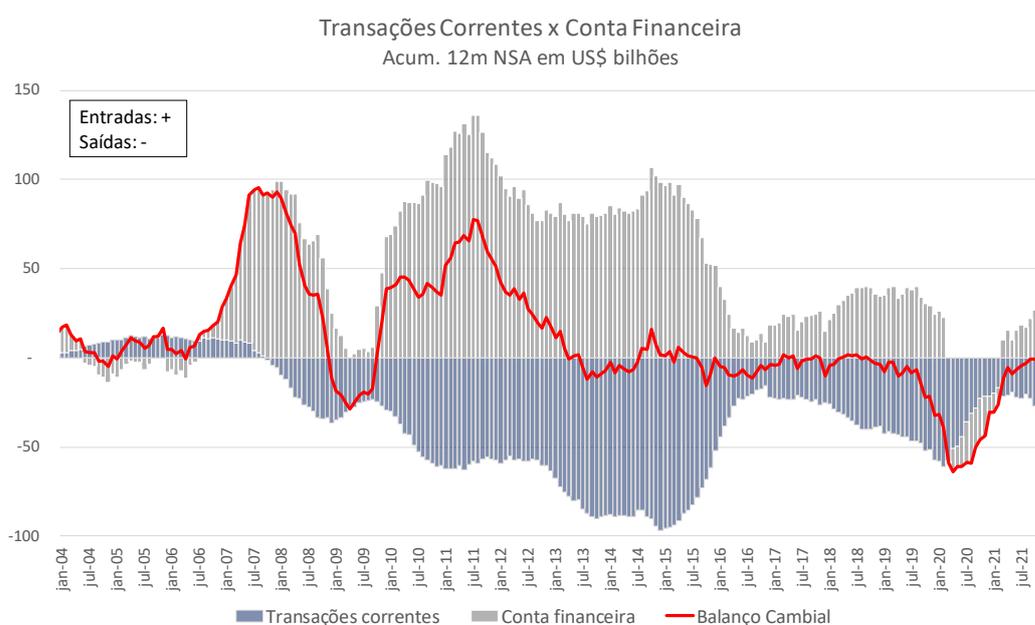


Figura 2 - Transações Correntes vs. Conta Financeira. Fonte: Banco Central do Brasil

Dessa forma, a contínua evasão de capitais, evidenciada na conta financeira, mais que compensava o ajuste nas contas externas, contribuindo para um déficit no mercado cambial (Figura 2) e manutenção do câmbio em patamares mais depreciados do que seus fundamentos poderiam sugerir.

Portanto, o estudo a respeito dos determinantes da conta financeira e, conseqüentemente, sua perspectiva futura, é imprescindível para a compreensão da dinâmica da taxa de câmbio.

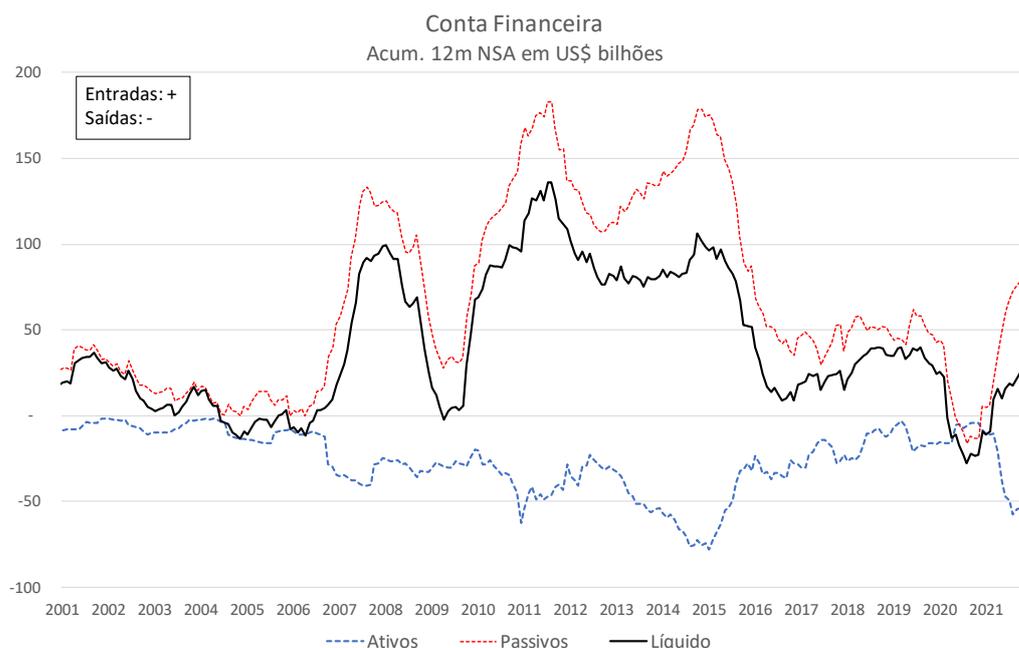


Figura 3 - Conta Financeira. Fonte: Banco Central do Brasil

A Figura 3 mostra os fluxos brutos de entrada (passivos), de saída (ativos) e o fluxo de capitais líquido. Curiosamente, as saídas brutas possuem um padrão semelhante às entradas brutas - isto é, quando não-residentes aumentam sua participação em ativos brasileiros, os residentes também aumentam sua participação em ativos externos - em linha com o observado por Ahmed e Zlate (2014) para os principais países emergentes asiáticos e latino-americanos. Apesar de sentidos que se compensam quando observados de forma agregada, os fluxos brutos de saída são, em geral, menores em magnitude que os fluxos brutos de entrada, de forma que o último apresente uma dinâmica similar ao resultado líquido. A escolha pelo fluxo bruto de entrada ou líquido como variável dependente é um importante ponto de debate na literatura, ao qual retornaremos nas seções posteriores.

A previsão dos fluxos de capitais não é uma tarefa fácil. Apesar disso, as evidências empíricas nos fornecem informações úteis a respeito da probabilidade de outros indicadores econômicos anteciparem tais fluxos. Dessa forma, o presente trabalho se insere na literatura propondo modelos de previsão um trimestre a frente dos fluxos de capitais líquidos para o Brasil. Para isso, utilizamos técnicas de *machine learning* e modelos de regularização para seleção das variáveis explicativas importantes e aferimos se tais modelos são capazes de gerar previsões fora da amostra melhores que um modelo AR(1), utilizado como *benchmark*. Os fluxos são obtidos a partir de dados trimestrais do balanço de pagamentos, englobando desde o primeiro trimestre de 2004 até o primeiro trimestre de 2021.

Os resultados apontam que os dois modelos sugeridos, LASSO e adaLASSO + OLS, aos quais retornaremos nas seções posteriores, geram previsões melhores que o modelo de *benchmark*. Apesar disso, quando comparados entre si, não podemos rejeitar a hipótese nula de que tais geram previsões estatisticamente iguais a um nível de significância de 10%.

O restante deste trabalho está organizado da seguinte forma: a seção 2 traz uma revisão dos trabalhos empíricos sobre os determinantes dos fluxos de capitais para países emergentes, a seção 3 descreve os dados utilizados, a seção 4 apresenta a estratégia empírica utilizada, a seção 5 apresenta e interpreta os principais resultados e a seção 6 conclui.

Revisão da literatura

Para uma melhor compreensão sobre o desenvolvimento da literatura até os dias de hoje, torna-se válido contextualizar, brevemente, o cenário macroeconômico no qual os trabalhos emergiram bem como as questões abordadas ao longo do tempo. Uma revisão mais abrangente dos estudos empíricos sobre fluxos de capitais para economias emergentes é fornecida por Koepke (2015) e Hannan (2018).

Com a retomada dos fluxos de capitais no início da década de 90 para países da América Latina, sob um cenário de recuperação pós-crise da dívida externa latino-americana e, simultaneamente, de recessão e taxa de juros em queda nos Estados Unidos (US), surgiram os primeiros questionamentos a respeito dos fluxos de capitais: eles refletiam a melhora dos fatores domésticos nos países de destino ou foram impulsionados por fatores externos que, provavelmente, seriam revertidos? Para endereçar essas questões, Calvo, Leiderman e Reinhart (1993) introduziram o arcabouço teórico que ficou popularmente conhecido como “*push vs. pull factors*” e que continua sendo utilizado até hoje de base para trabalhos empíricos. Em resumo, “*push factors*” são fatores externos (ou fatores sob a ótica da oferta) que sustentam a oferta de liquidez global como, por exemplo, aversão ao risco global, preço de commodities, PIB US e taxa de juros US. “*Pull factors*”, por outro lado, são fatores domésticos (ou fatores sob a ótica da demanda) que influenciam risco e retorno aos investidores como, por exemplo, fundamentos macroeconômicos, políticas domésticas e componentes cíclicos como PIB e taxa de juros. Embora reconheçam que a melhora dos fatores domésticos tenha sido importante, os resultados obtidos apontam que os fatores externos, em especial a taxa de juros dos Estados Unidos, foram os determinantes para a retomada dos fluxos. Em trabalho similar, Fernandez-Arias (1996) obteve resultados semelhantes.

Com o advento da crise financeira global de 2008/09, que provocou uma grande retração dos fluxos de capitais no mundo inteiro, a literatura continuou buscando entender a importância relativa dos “*push vs. pull factors*”, concentrando-se nos determinantes de episódios extremos.

Nessa linha, o aumento da granularidade dos dados levantou um importante ponto de debate: a utilização como variável dependente do fluxo bruto de entrada de capital (considerando apenas o fluxo de não-residentes) ou do fluxo de capital líquido (incorporando residentes). Forbes e Warnock (2012) analisaram mais de 50 países emergentes e desenvolvidos e mostraram que os determinantes do primeiro podem ser consideravelmente diferentes dos do segundo – que é mais utilizado na literatura. Os autores argumentam que, no início da década de 90, o fluxo líquido era, aproximadamente, reflexo do fluxo bruto de entrada (isto é, de não-residentes), o que justificava a interpretação de que o segundo determinava o primeiro. Por outro lado, no período mais recente da janela de análise, que vai de 1980 até 2009, constatam um aumento do volume e da volatilidade nos fluxos brutos (tanto de não-

residentes quanto de residentes) em contraste com certa estabilidade no fluxo líquido, tornando importante a diferenciação entre eles uma vez que agentes domésticos e externos podem responder a choques de formas distintas, compensando (conforme observado pelos autores) ou amplificando o resultado líquido. O estudo desenvolve um *framework* para analisar episódios de fluxos extremos, classificando-os em quatro categorias: “*surges*”/“*stops*” (aumento/queda acentuado(a) nas entradas brutas por não-residentes) e “*flight*”/“*retrenchment*” (aumento/queda acentuado(a) nas saídas brutas por residentes). Por fim, concluem que fatores globais, notadamente aversão ao risco global, são os principais determinantes de tais episódios enquanto fatores domésticos são menos importantes.

Em estudo similar, porém, com foco apenas em países emergentes (*EMEs – Emerging Market Economies*), Ghosh et. al (2014) analisaram as causas de acentuados aumentos no fluxo de capital líquido, classificados como “*surges*”, para 56 EMEs no período de 1980-2011. Em referência ao trabalho de Forbes e Warnock (2012), diferenciam tais episódios entre “*liability-driven*” (impulsionados por não-residentes) e “*asset-driven*” (impulsionados por residentes), reconhecendo que cada tipo pode ter diferentes propriedades e implicações. Apesar disso, justificam a escolha pelo fluxo líquido na identificação dos eventos para garantir que eles correspondessem a períodos extremos de entrada de capital, o que não é necessariamente verdade no *framework* de Forbes e Warnock (2012). Concluem que fatores globais (“*push factors*”) - como a taxa de juros US e aversão ao risco global - atuam como, conforme definição dos autores, “*gatekeepers*” que determinam a frequência dos eventos enquanto fatores domésticos (“*pull factors*”) - como necessidade de financiamento externo, abertura da conta de capital e regime da taxa de câmbio - são importantes determinantes de sua magnitude.

Conceitualmente, a escolha da variável dependente deveria decorrer da questão abordada pelo trabalho. O fluxo líquido pode ser mais relevante para a apreciação/depreciação da taxa de câmbio e preocupações com o superaquecimento da economia, enquanto os fluxos brutos podem ser mais relevantes para questões de estabilidade financeira e da capacidade dos sistemas financeiros em intermediar efetivamente os fluxos (Ahmed e Zlate, 2014).

Boa parte da literatura a respeito dos determinantes dos fluxos de capitais para EMEs optou pela ênfase aos episódios de fluxos extremos sob a premissa de que tais episódios possuíam determinantes diferentes dos períodos “normais”. Apesar de ser uma estratégia interessante, a principal desvantagem é a de não nos permitir mensurar o quanto desses fluxos é resultado de movimentos nas variáveis explicativas que influenciam tais fluxos em modelos que englobam períodos “normais” em sua janela amostral.

Sob essa ótica e, diferentemente dos estudos apresentados nesta seção, Ahmed e Zlate (2014) examinaram os determinantes dos fluxos de capitais para 12 países emergentes entre 2002 e 2012 sem a tradicional abordagem de “*push vs. pull factors*”. Ao invés disso, mostraram que, além da usual aversão ao risco global, os diferenciais de crescimento e de taxa de juros entre os países emergentes e

desenvolvidos eram estatisticamente significativos e importantes para o fluxo líquido de capitais. Além disso, utilizando a crise financeira global de 2008 para dividir a amostra, identificaram maior sensibilidade dos fluxos ao diferencial de taxa de juros e à aversão ao risco no período pós crise, especialmente os de investimento em portfólio.

Por fim, se concentrando na desaceleração dos fluxos de capitais e na evolução do cenário após a crise financeira global, o trabalho produzido pelo IMF (2016), utilizando dados de 2011 a 2016, verificou que a desaceleração foi resultado tanto de um fluxo bruto de entrada mais fraco quanto de um fluxo bruto de saída mais forte, explicados pela redução no diferencial de crescimento entre países emergentes e desenvolvidos. Em trabalho similar, Hannan (2017) conclui que a desaceleração observada nos fluxos foi decorrente, sobretudo, da combinação entre menor crescimento nos países destinatários e piora no sentimento de risco global.

O presente trabalho se diferencia dos apresentados nesta seção sobretudo no que diz respeito ao objetivo. Nesse sentido, em vez de forçarmos no entendimento da importância relativa das covariadas na dinâmica do fluxo já observado, estamos interessados no que podemos inferir sobre sua dinâmica futura. Dessa forma, utilizando variáveis apontadas pela literatura como precursores do fluxo de capital, realizamos previsões um trimestre a frente nos utilizando de técnicas de *machine learning* e modelos de regularização para a seleção das variáveis importantes. Outro ponto a ser destacado é que, conforme observado por NG (2021), a pandemia COVID-19 é um evento global de proporções históricas que, em poucos meses, mudou as propriedades das séries temporais de forma que muitos modelos de previsão pré-covid se tornaram inadequados. Sob essa perspectiva, o presente trabalho acrescenta a literatura introduzindo indicadores da pandemia numa tentativa de limpar os dados dos relevantes e persistentes choques exógenos que provocaram alterações na dinâmica dos fluxos na margem.

Dados

O estudo empírico que será apresentado a seguir tem como foco os fluxos financeiros em termos líquidos em que há liquidação financeira no mercado de câmbio (e, portanto, causam impacto cambial) no período entre o primeiro trimestre de 2004 até o primeiro trimestre de 2021. Conforme explicado anteriormente, tais fluxos foram obtidos através do Balanço de Pagamentos pela soma dos quatro grupos que compõe a conta financeira: investimentos diretos (exceto reinvestimento de lucros), investimentos em carteira (exceto reinvestimento de juros), derivativos e outros investimentos.

Em vez de analisarmos cada tipo individualmente e/ou, como em outros estudos, classificarmos segundo sua origem (ativo/passivo), o objetivo principal desse trabalho é fornecer um modelo que proporcione uma boa estimativa para o fluxo de capitais de forma consolidada, justificando a escolha pelo fluxo líquido total como variável de interesse. Além disso, outro ponto que contribui para essa escolha é o aumento da complexidade para distinção da natureza dos fluxos, não sendo incomuns revisões nas séries do Balanço de Pagamentos para reclassificá-las, o que poderia prejudicar a estimação caso cada grupo fosse modelado individualmente.

Em relação às variáveis explicativas, tal como em diversos trabalhos aqui mencionados, incluímos no modelo variáveis que se enquadram no framework de “*push vs. pull factors*” e variáveis que são diferenciais vis-à-vis aos Estados Unidos (diferenciais da taxa de juros e de crescimento).

Para refletir as expectativas dos agentes a respeito do dinamismo da economia brasileira frente o resto do mundo, utilizamos o diferencial de crescimento medido pela diferença entre a taxa de crescimento real trimestral dessazonalizado do Brasil e dos Estados Unidos, utilizada como proxy para a taxa de crescimento de países desenvolvidos.

Na mesma linha, representando a possibilidade de rendimento no Brasil vis-à-vis o restante do mundo, utilizamos a variação em pontos percentuais do diferencial da taxa de juros, medido pela diferença entre a taxa DI brasileira e a taxa de juros dos Estados Unidos (*Federal Fund Rate*), utilizada como proxy da taxa de juros de referência externa e, portanto, a alternativa de rendimento que os agentes dispõem em países desenvolvidos.

De acordo com Froot et. Al (2001) em estudo para 44 países, rendimentos passados foram significativos para explicar a entrada de recursos para o mercado de ações, evidenciando a importância do desempenho dos índices de ações para as decisões de investimentos estrangeiros, que procuram engajar-se na tendência dos mercados. Dessa forma, buscando refletir essas expectativas de rendimento dos agentes em relação ao mercado acionário brasileiro, utilizamos a taxa de crescimento do índice Bovespa.

Como indicadores de aversão ao risco, optamos pela variação do *Chicago Board Options Exchange Market Volatility Index* (VIX), indicador de volatilidade do índice S&P500 e proxy para a aversão ao risco global. Além disso, também utilizamos a variação do *CDS* (*Credit Default Swap*) brasileiro, um título derivativo que funciona como um seguro contra a inadimplência em operações de crédito, como proxy para o risco-país englobando os fundamentos da política fiscal.

Sendo um país com participação de *commodities* superior a 60% na pauta exportadora e, dado o co-movimento entre os preços de *commodities* e o câmbio dos países exportadores (*commodity currencies*), utilizamos a taxa de crescimento dos termos de troca como uma possibilidade/expectativa de rendimento dos agentes internacionais.

Por fim, na tentativa de capturar os choques exógenos causados pelo COVID-19, que se mostram relevantes e persistentes o suficiente para provocar alterações na dinâmica dos fluxos no período da amostra, utilizamos o *Google Mobility Report*, que mostra tendências de movimento em diferentes categorias de lugares (varejo e recreação; mercearia e farmácia; parques; estações de transporte; locais de trabalho; residencial). Cabe destacar que, em vez de utilizarmos a média simples dos 6 índices, retiramos as categorias residencial e parques para colocar mais peso no índice para a atividade econômica.

Os símbolos das variáveis, sua descrição e fonte são apresentados na Tabela 1. A Tabela 2 resume as estatísticas descritivas.

Tabela 1 - Dados

Variável	Descrição	Fonte	Periodicidade
<i>cds</i>	<i>Brasil CDS USD SR 5YD14</i>	CBIN	Diária
<i>juros_BR</i>	<i>Brazil Cetip DI Interbank Deposit Rate</i>	CETIP	Diária
<i>juros_US</i>	<i>Federal Funds Effective Rate</i>	<i>Federal Reserve Bank of New York</i>	Mensal
<i>ibov</i>	Índice Bovespa	BBG	Mensal
<i>Ks</i>	Conta Financeira ex. Erros e Omissões (Hiato Financeiro)	BCB	Mensal
<i>mobility</i>	<i>Google Mobility Brasil Ajustado</i>	<i>Google</i>	Diária
<i>pibqoq_BR</i>	<i>Brazil GDP QoQ SA 1995=100</i>	IBGE	Trimestral
<i>pibqoq_US</i>	<i>GDP US Chained 2012 Dollars QoQ Unannualized</i>	<i>U.S. Bureau of Economic Analysis</i>	Trimestral
<i>tot</i>	Termos de Troca	Ministério da Indústria, Comércio Exterior e Serviços	Mensal
<i>vix</i>	<i>Chicago Board Options Exchange Volatility Index</i>	BBG	Diária

Tabela 2 - Estatísticas descritivas

Variável	Obs.	Média	Desv. Padrão	Mín.	Máx.
<i>cds</i>	69	-45	186	-1154	271
<i>ibov</i>	69	3.1%	10.3%	-32.4%	28.3%
<i>juros_dif</i>	69	-0.16%	1.0%	-2.8%	2.4%
<i>Ks</i>	69	11,943	13,385	-28,936	46,281
<i>mobility</i> ¹	5	-15.7%	11.5%	-34.5%	-3.7%
<i>pib_dif</i>	69	0.1%	1.2%	-3.2%	2.0%
<i>tot</i>	69	0.6%	3.4%	-6.9%	8.8%
<i>vix</i>	69	0.3	9.9	-35.5	-36.6

Notas: ¹ Média e desvio padrão calculados para os valores a partir do primeiro trimestre de 2020, quando a variável *mobility* assume valores diferentes de 0.

Estratégia empírica

Esta seção descreve os métodos usados neste trabalho para modelar e fazer previsões um trimestre à frente do fluxo de capital líquido usando dados trimestrais desde o primeiro trimestre de 2004 até o primeiro trimestre de 2021, último período com taxas de crescimento já divulgadas, totalizando 69 observações. Uma análise preliminar da Figura 4 nos permite observar uma boa correlação em nível entre os fluxos e o CDS, sobretudo a partir do último trimestre de 2005. Dessa forma, a data de início em 2004 nos possibilita filtrar da janela amostral um período marcado por um CDS explosivo e com baixa correlação com os fluxos, o que poderia prejudicar a estimação.

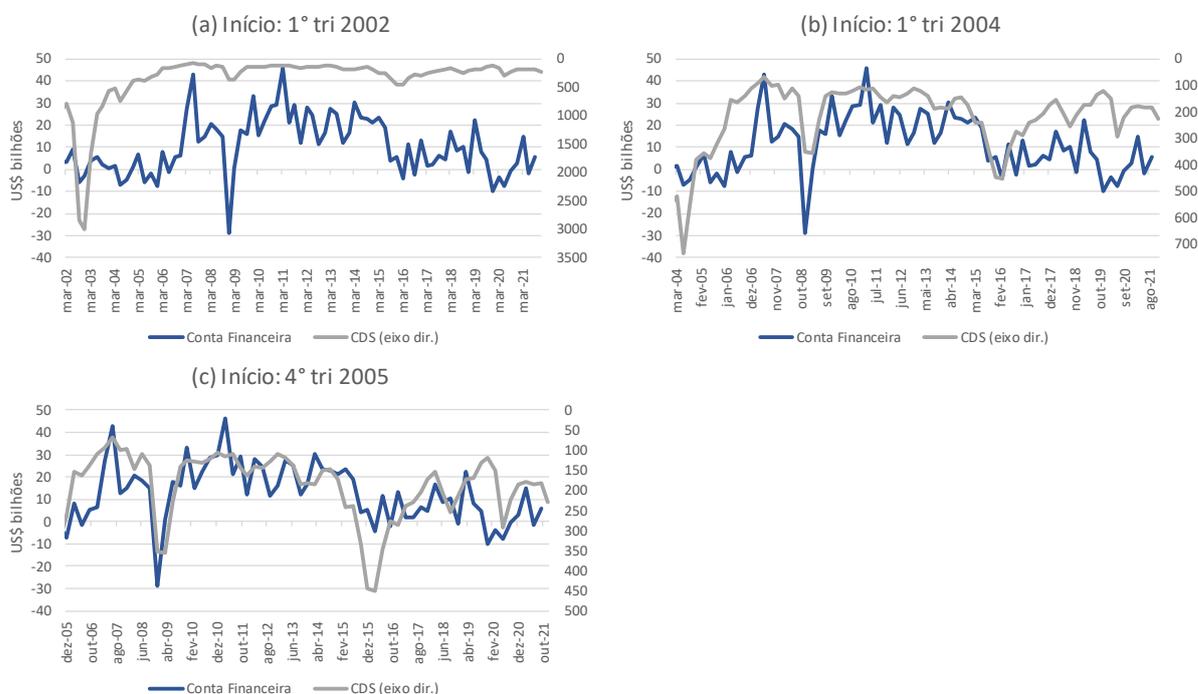


Figura 4 - Conta Financeira em US\$ bilhões vs. CDS (eixo direito, invertido)

Cabe ressaltar que a escolha do tamanho da janela amostral envolve um equilíbrio entre fatores opostos. Por um lado, uma janela mais curta implica um conjunto de dados menor para realização das previsões. Por outro lado, uma janela mais longa pode aumentar a probabilidade de ter havido mudanças no processo de geração dos dados ao longo do período analisado, de forma que dados antigos não sejam mais representativos para a dinâmica atual.

Nos trabalhos mencionados na revisão de literatura, o objetivo central era entender a importância relativa das covariadas na dinâmica do fluxo de capitais observada, justificando a utilização de uma janela amostral única para a estimação do modelo final. Por outro lado, no presente trabalho, estamos interessados no que podemos inferir sobre a dinâmica futura dos fluxos. Portanto, nos diferenciamos dos demais utilizando uma proposta de janela crescente, conforme ilustrado na

Figura 5. Dessa forma, a cada passo de estimação e previsão, a data fim do modelo é deslocada em um trimestre. Escolhemos arbitrariamente o quarto trimestre de 2017 como data fim do primeiro passo, totalizando 13 previsões um trimestre à frente.

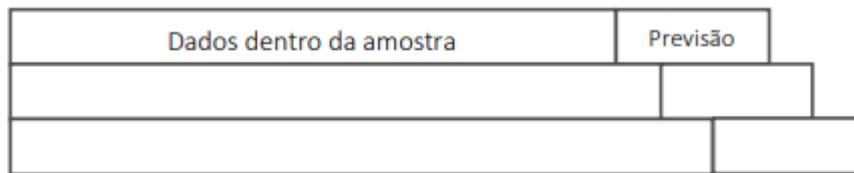


Figura 5 - Janela crescente ou rolling forward

Para evitar problemas associados a não estacionariedade das séries, aplicamos o teste de Dickey-Fuller Aumentado (ADF). Com isso, rejeitamos a hipótese nula de presença de raiz unitária a um nível de significância de 10%, concluindo que todas as séries utilizadas são $I(0)$ e, portanto, estacionárias em nível.

Consideramos uma abordagem de previsão onde o fluxo de capitais líquido um trimestre a frente, Ks_{t+1} , é modelado em função de um conjunto de preditores medidos no tempo t :

$$Ks_{t+1} = \beta'x_t + \varepsilon_{t+1}$$

Onde ε_{t+1} é o erro de previsão, $x_t = (x_{1t}, \dots, x_{qt})$ é o conjunto de q preditores disponíveis em t , no qual podem estar incluídos valores defasados do fluxo de capitais líquido, e $\beta \in \mathbb{R}^q$ é o vetor de parâmetros a ser estimado. Conforme discutido na seção anterior, os preditores disponíveis em t são valores defasados ou não das seguintes variáveis:

$$x_t = \begin{pmatrix} cds \\ ibov \\ juros_dif \\ Ks \\ mobility \\ pib_dif \\ tot \\ vix \end{pmatrix}$$

Onde Ks são valores defasados do fluxo de capitais líquido, pib_dif e $juros_dif$ representam os diferenciais de crescimento e de taxa de juros, $ibov$ é o índice Bovespa, vix é o indicador de aversão ao risco global, cds é a proxy para o risco-país, tot são os termos de troca e $mobility$ é o indicador de mobilidade durante a pandemia.

Modelo empírico

Muitas vezes, ao utilizarmos o processo de otimização de mínimos quadrados ordinários (ou *OLS – Ordinary Least Squares*) para encontrar o melhor ajuste para um conjunto de dados, utilizamos variáveis explicativas que podem não estar realmente associadas à variável dependente. É difícil que esse processo de

otimização force, naturalmente, o coeficiente dessas variáveis a zero. Dessa forma, além da possibilidade de obtermos um modelo excessivamente ajustado aos dados (*overfitting*), podemos aumentar, desnecessariamente, sua complexidade e prejudicar sua interpretabilidade.

Com o objetivo de evitar esses problemas e, assim, melhorar a qualidade da predição, utilizamos modelos de regularização. A ideia por trás é penalizar os parâmetros que correspondem a variáveis irrelevantes, levando-os a zero. Dentre esses modelos, o LASSO (*least absolute shrinkage and selection operator*), introduzido por Tibshirani (1996), e o adaLASSO (*adaptive LASSO*) de Zou (2006) receberam atenção especial na literatura. Apesar de não ser o caso no presente trabalho, foi mostrado que o LASSO pode lidar com mais variáveis do que observações, selecionando corretamente o subconjunto de variáveis relevantes (Zhao e Yu, 2006). Apesar disso, conforme observado por Zhao e Yu (2006) e Zou (2006), o LASSO requer uma forte condição denotada “condição irrepresentável” para alcançar a consistência na seleção do modelo, além de não ter a propriedade oráculo. Para contornar essas deficiências, Zou (2006) propôs o adaLASSO, que é uma metodologia de dois estágios que usa um estimador, geralmente o LASSO, para ponderar a importância relativa dos regressores.

O estimador LASSO é definido por:

$$\hat{\beta} = \arg \min_{\beta} \left[\sum_{t=1}^T (Ks_{t+1} - \beta'x_t)^2 + \lambda \sum_{j=1}^q |\beta_j| \right]$$

Onde λ é o parâmetro de penalização e é determinado por técnicas orientadas por dados (*data-driven techniques*), como validação cruzada (*cross-validation*) ou o uso de critérios de informação (*information criteria*). Medeiros e Mendes (2016) sugerem que a seleção de λ utilizando o critério de informação Bayesiano (BIC) proporciona resultados melhores, tanto em termos de precisão quanto em tempo de computação, do que os obtidos caso a seleção fosse feita através da validação cruzada, especialmente no caso de séries temporais, onde a utilização do último método é mais complicada. Dessa forma, selecionamos o parâmetro de penalização via BIC, tanto no LASSO quanto no adaLASSO.

O adaLASSO é definido por:

$$\hat{\beta} = \arg \min_{\beta} \left[\sum_{t=1}^T (Ks_{t+1} - \beta'x_t)^2 + \lambda \sum_{j=1}^q \omega_j |\beta_j| \right]$$

Onde $\omega_j = |\hat{\beta}_j^*|^{-\tau}$ representa diferentes pesos a penalização de cada variável, $\hat{\beta}_j^*$ é o parâmetro estimado no primeiro estágio, e $\tau > 0$ determina o quanto queremos enfatizar a diferença nos pesos. Estimamos o adaLASSO usando o LASSO como primeiro estágio e $\tau = 1$, que é o valor mais utilizado na literatura.

Belloni e Chernozhukov (2013) mostraram que estimar uma regressão linear usando as variáveis selecionadas pelo LASSO (pós-OLS) funciona pelo menos tão

bem quanto usar apenas o próprio LASSO em termos de taxa de convergência para o oráculo, além de ter um viés menor. Portanto, estimamos a regressão pós-OLS para o adaLASSO (adaLASSO + OLS) na tentativa de obtermos menores erros de previsão.

Resultados

Para a escolha do modelo a ser apresentado, adotamos algumas estratégias. Conforme abordado na seção anterior, a utilização do primeiro trimestre de 2004 como possível data de início da janela amostral se justifica pela tentativa de filtrar um período marcado por um CDS explosivo, onde a correlação entre essa variável e os fluxos não aparentava ser representativa para a dinâmica atual, como pode ser observado na Figura 6. Além disso, com o objetivo de minimizar a perda do conjunto de dados para a realização das previsões, limitamos a data de início da janela amostral ao último trimestre de 2005, totalizando 8 possíveis datas de início.

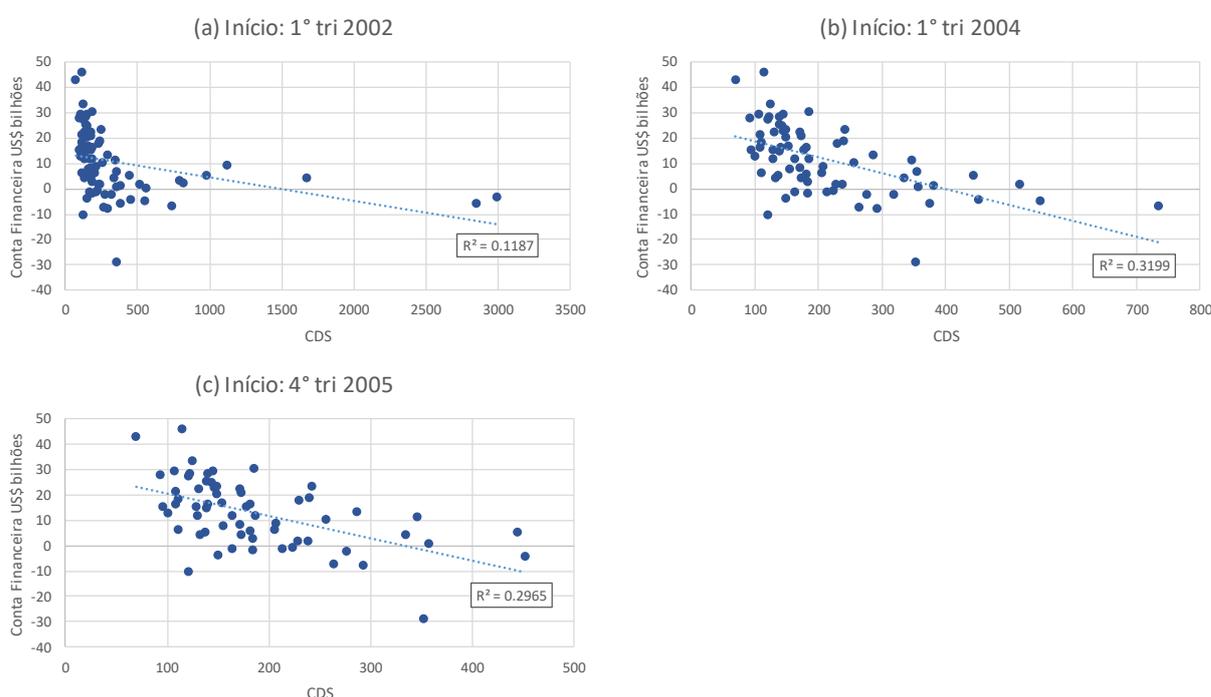


Figura 6 - Dispersão Conta Financeira vs. CDS

Para cada data de início, permitimos a inclusão ou não da variável *mobility* e a utilização de até 3 defasagens (de $t0$ até $t-3$) de todas as covariadas entre as variáveis explicativas, totalizando 64 possíveis conjuntos de dados para a realização das estimações.

Escolhemos arbitrariamente o quarto trimestre de 2017 como data final do primeiro passo para a estimação do modelo dentro da amostra (*in sample*) e, em seguida, realizamos a previsão um trimestre a frente (*out of sample*) em todos os possíveis conjuntos. A cada passo, utilizamos a estratégia de janela crescente, isto é, mantivemos a data inicial fixa e deslocamos a data final e a previsão em um trimestre, totalizando 13 passos até a previsão do último fluxo, referente ao primeiro trimestre de 2021.

No Anexo 1, são apresentados os erros de previsão dos modelos (LASSO e adaLASSO + OLS) em cada um dos 64 possíveis conjuntos, além dos erros de

previsão para um modelo AR(1), utilizado como *benchmark*. O modelo com menor erro de previsão para cada conjunto é exibido em negrito. Como métricas de erro, optamos pela raiz do erro quadrático médio (*RMSE* - *root mean squared error*) e pelo erro médio absoluto (*MAE* - *mean absolut error*), definidos por:

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2}$$

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j|$$

Onde $n = 13$ é o número total de passos.

A Tabela 3 apresenta a média da raiz do erro quadrático médio e do erro médio absoluto de todas as possíveis datas de início para os conjuntos estimados. Uma análise preliminar nos permite inferir que a variável *mobility*, introduzida como uma tentativa de capturar os choques exógenos causados pelo COVID-19, não foi selecionada pelo modelo *adaLASSO* + *OLS* na maior parte dos conjuntos estimados, fato observado pela ausência de diferença nas métricas de erro para esse modelo nos conjuntos estimados com e sem a inclusão dessa variável. O modelo *LASSO*, por outro lado, apesar de selecionar a variável *mobility* com maior frequência entre as variáveis importantes, sobretudo nos conjuntos nos quais foi permitida a inclusão de 1 e 2 defasagens de todas as covariadas, apresentou uma piora na capacidade preditiva em comparação aos conjuntos nos quais a inclusão dessa variável não foi permitida.

Além disso, o Anexo 1 mostra que, em 63 dos 64 conjuntos estimados, os modelos *LASSO* e/ou *adaLASSO* + *OLS* obtiveram erros menores que o *benchmark*. Os resultados da Tabela 3 sugerem que a melhora na capacidade preditiva se deu, sobretudo, nos conjuntos onde foram permitidas a inclusão de até 2 defasagens de todas as covariadas entre as variáveis explicativas. Nesses conjuntos, os resultados apontam para uma redução de, em média, 17% e 20% nos erros de previsão (raiz do erro quadrático médio) para os modelos *LASSO* e *adaLASSO* + *OLS*, respectivamente, em relação ao *benchmark*.

Tabela 3 - Média da raiz do erro quadrático médio e do erro médio absoluto

<i>mobility</i>	Erro Médio para todas as datas de início		Aumento/Redução do erro em relação ao <i>benchmark</i>	
	Não	Sim	Não	Sim
Sem lags				
LASSO	9,994 (8,406)	9,994 (8,406)	-2.4% -(3.2%)	-2.4% -(3.2%)
adaLASSO + OLS	9,612 (7,412)	9,612 (7,412)	-6.1% -(14.6%)	-6.1% -(14.6%)
Até t-1				
LASSO	9,855 (7,876)	10,070 (8,055)	-3.8% -(9.3%)	-1.7% -(7.2%)
adaLASSO + OLS	10,156 (8,050)	10,156 (8,050)	-0.8% -(7.3%)	-0.8% -(7.3%)
Até t-2				
LASSO	8,502 (5,821)	9,074 (6,308)	-17.0% -(33.0%)	-11.4% -(27.4%)
adaLASSO + OLS	8,168 (5,537)	8,168 (5,537)	-20.2% -(36.2%)	-20.2% -(36.2%)
Até t-3				
LASSO	9,342 (7,336)	9,342 (7,335)	-8.8% -(15.5%)	-8.8% -(15.5%)
adaLASSO + OLS	9,429 (7,036)	9,431 (7,051)	-7.9% -(19.0%)	-7.9% -(18.8%)
Benchmark				
AR(1)	10,240 (8,683)			

Notas: As colunas 2 e 3 mostram a média da raiz do erro quadrático médio e o erro médio absoluto (entre parênteses) de todas as possíveis datas de início. As colunas 4 e 5 mostram o aumento/redução da raiz do erro quadrático médio e do erro médio absoluto (entre parênteses) em relação ao *benchmark*

Novamente, como estamos interessados no que podemos inferir sobre a dinâmica futura dos fluxos, optamos por discutir os resultados do conjunto com o menor erro de previsão fora da amostra, obtido pela menor média da raiz do erro quadrático médio dos modelos LASSO e adaLASSO + OLS dentre os 64 conjuntos estimados. No conjunto em questão, a data de início da janela de estimação refere-se ao primeiro trimestre de 2004, não permitimos a inclusão da variável *mobility* e permitimos até 2 defasagens (de t_0 até $t-2$) de todas as covariadas entre as variáveis explicativas. Os erros dos modelos (observado – estimado) deste conjunto para os 13 passos de previsão fora da amostra são apresentados na Figura 7.

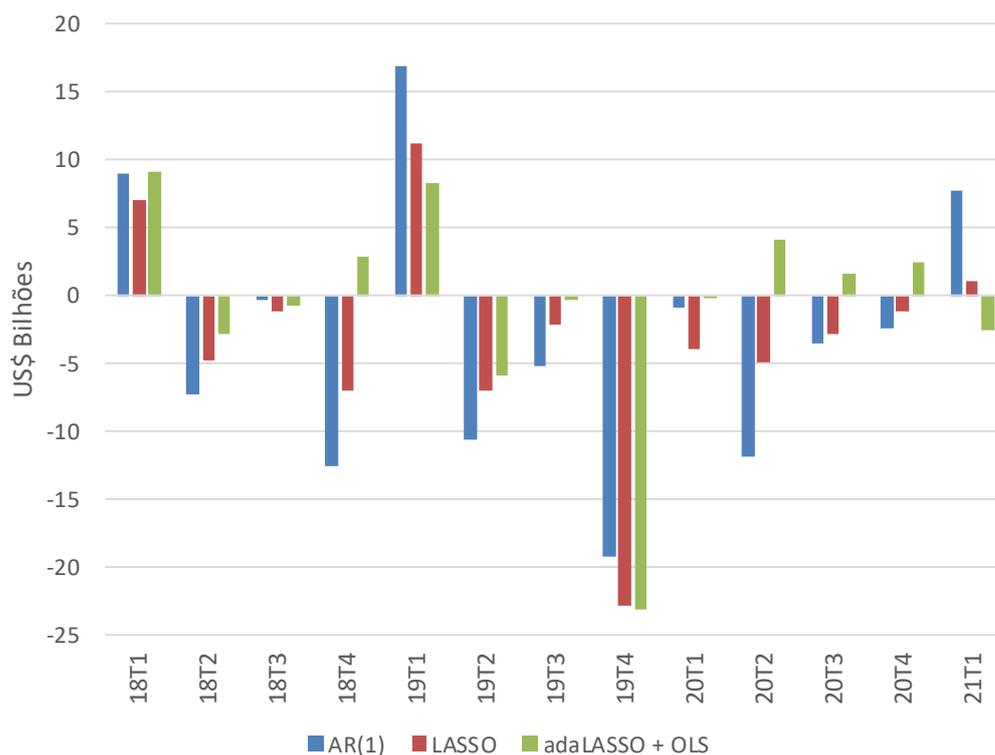


Figura 7 - Erros de previsão dos modelos fora da amostra em US\$ bilhões

Na Figura 7, é possível observar um erro de previsão acima da média no último trimestre de 2019, período no qual todos os modelos superestimaram o fluxo de capital líquido. Uma possível explicação é a de que de tal período engloba o início dos casos da pandemia COVID-19 na China, sendo marcado por uma forte evasão de capitais concentrada no mês de dezembro. Dessa forma, é natural que os preditores disponíveis até o terceiro trimestre de 2019 não fossem capazes de antecipar tamanha retração no fluxo de capitais.

A forma usual de avaliar o desempenho de previsão entre dois modelos concorrentes é aplicar o teste proposto por Diebold-Mariano (1995), que avalia se as previsões fora da amostra de dois modelos são estatisticamente iguais. Dessa forma, os modelos propostos foram testados contra o *benchmark* e entre si. A Tabela 4 resume os resultados.

Tabela 4 - P-valor dos testes Diebold-Mariano

		Modelo 1	
		LASSO	adaLASSO + OLS
Modelo 2	LASSO	X	0.1216
	adaLASSO + OLS	0.8784	X
	AR (1)	0.0699	0.0673

Notas: H_0 Os dois modelos possuem a mesma precisão de previsão. H_1 O modelo 2 é menos preciso que o modelo 1

A um nível de significância de 10%, rejeitamos a hipótese nula de que os modelos propostos possuem a mesma precisão de previsão do modelo de *benchmark*. Com isso, aceitamos a hipótese alternativa de que os dois modelos propostos geram previsões fora da amostra mais precisas que o modelo de *benchmark*. Em relação à raiz do erro quadrático médio, os resultados apresentados na Tabela 5 apontam para uma redução de 18% e 23% nos erros de previsão para os modelos LASSO e adaLASSO + OLS, respectivamente, em comparação ao *benchmark*. Já em relação ao erro médio absoluto, tal redução foi de 28% e 40%, respectivamente.

Por fim, ainda que o modelo adaLASSO + OLS tenha obtido erros menores que o modelo LASSO, os resultados apresentados na Tabela 4 não nos permitem rejeitar, a um nível de significância de 10%, a hipótese nula de que esses modelos possuem a mesma precisão de previsão sendo, portanto, estatisticamente iguais.

Tabela 5 - Raiz do erro quadrático médio e erro médio absoluto do conjunto selecionado

	LASSO	adaLASSO + OLS	AR (1)
<i>RMSE</i>	8,207	7,709	10,032
<i>MAE</i>	5,944	4,927	8,276

Notas: Os valores em negrito representam o melhor modelo de acordo com cada medida de erro

Conclusão

A previsão dos fluxos de capitais não é uma tarefa fácil. Nesse trabalho, comparamos diferentes métodos de previsão aplicados aos dados de fluxo de capital líquido brasileiros obtidos a partir de dados trimestrais do balanço de pagamentos.

Utilizando variáveis econômicas consistentes com as evidências apresentadas nos estudos abordados na Revisão da Literatura como potenciais precursores do fluxo de capitais, os modelos propostos, tanto LASSO quanto adaLASSO + OLS, foram capazes de gerar previsões fora da amostra melhores que o modelo de *benchmark*, AR (1), a um nível de significância de 10%.

Esses resultados são os mesmos independentemente da métrica de erro utilizada, seja raiz do erro quadrático médio ou erro médio absoluto. Em relação à primeira, os resultados apontam para uma redução de 18% e 23% nos erros de previsão para os modelos LASSO e adaLASSO + OLS, respectivamente, em comparação ao *benchmark*. Já em relação à segunda, tal redução foi de 28% e 40%, respectivamente.

Apesar disso, quando comparados entre si, não podemos descartar a hipótese nula de que os modelos propostos geram previsões fora da amostra estatisticamente iguais, embora o modelo adaLASSO + OLS tenha obtido erros menores.

Por fim, a variável *mobility*, introduzida como uma tentativa de capturar os choques exógenos causados pelo COVID-19, não foi selecionada na maioria dos conjuntos nos quais permitimos sua inclusão entre as variáveis explicativas. Além disso, nos conjuntos onde ela foi selecionada, observamos uma piora nas métricas de erro em comparação aos conjuntos nos quais a inclusão dessa variável não foi permitida.

Referências

Ahmed, S., and A. Zlate, 2014, “Capital Flows to Emerging Market Economies: A Brave New World?”, *Journal of International Money and Finance*, Vol. 48, pp. 221–248.

Belloni, A., and Chernozhukov, V., 2013, “Least squares after model selection in high-dimensional sparse models.”, *Bernoulli*, 19(2), 521–547.

BPM6, 2011, “Balance of payments and international investment position manual.”, Disponível em:
<<https://www.imf.org/external/pubs/ft/bop/2007/pdf/bpm6.pdf>>

Calvo, G.A., L. Leiderman, and C.M. Reinhart, 1993, “Capital Inflows and Real Exchange Rate Appreciation in Latin America: the Role of External Factors.”, *IMF Staff Papers*: 108–151.

Diebold, F. and Mariano, R., 1995, “Comparing Predictive Accuracy.”, *Journal of Business and Economic Statistics*, 13(3), 253-263.

Fernandez-Arias, E., 1996, “The New Wave of Private Capital Inflows: Push or Pull?”, *Journal of Development Economics*, Vol. 48, No. 2, pp. 389–418.

Forbes, K.J., and F.E. Warnock, 2012, “Capital Flow Waves: Surges, Stops, Flight, and Retrenchment”, *Journal of International Economics*, Vol. 88, No. 2, pp. 235–51.

Froot, K. A., O’Connell, P. G., and Seasholes, M. S., 2001, “The portfolio flows of international investors”, *Journal of Financial Economics*, Vol. 59, No. 2, pp. 151–193.

Garcia, M. G., Medeiros, M. C., and Vasconcelos, G. F., 2017, “Real-time inflation forecasting with high-dimensional models: The case of Brazil”, *International Journal of Forecasting*, 33(3), 679-693.

Ghosh, A.R., M. S. Qureshi, J. Kim, and J. Zalduendo, 2014, “Surges”, *Journal of International Economics*, Vol. 92, No. 2, pp. 266–85.

Hannan, S.A., 2017, “The Drivers of Capital Flows in Emerging Markets Post Global Financial Crisis,” *Journal of International Commerce, Economics and Policy*, Vol. 8, No. 2, pp. 1–28.

Hannan, S.A., 2018, “Revisiting the determinants of capital flows to emerging markets – A survey of the evolving literature”, *IMF Working Paper* no. 18/214.

IMF, 2016, “Understanding the Slowdown in Capital Flows in Emerging Markets,” *World Economic Outlook* (Washington, D.C.: International Monetary Fund) Chapter 2, April.

Koepke, R., 2015, “What Drives Capital Flows to Emerging Markets? A Survey of the Empirical Literature”, Institute of International Finance.

Medeiros, M., and Mendes, E., 2016, “ ℓ_1 -regularization of high-dimensional time-series models with flexible innovations.”, *Journal of Econometrics*, 191, 255–271.

NG, S., 2021, “Modeling macroeconomic variations after COVID-19.”, National Bureau of Economic Research.

Tibshirani, R., 1996, “Regression shrinkage and selection via the lasso.”, *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.

Zhao, P., and Yu, B., 2006, “On model selection consistency of LASSO.”, *The Journal of Machine Learning Research*, 7, 2541–2563.

Zou, H., 2006, “The adaptive LASSO and its oracle properties.”, *Journal of the American Statistical Association*, 101, 1418–1429.

Anexo 1 – Conjuntos estimados

(continua)

Início da Janela Amostral	2004/T1	2004/T2	2004/T3	2004/T4	2005/T1	2005/T2	2005/T3	2005/T4
Sem variável <i>mobility</i>								
Sem lags								
LASSO	10,085 (8,636)	9,798 (8,308)	9,882 (8,261)	9,925 (8,317)	9,955 (8,345)	10,059 (8,492)	10,067 (8,386)	10,177 (8,506)
adaLASSO + OLS	10,010 (8,219)	9,332 (7,129)	9,600 (7,374)	9,531 (7,262)	9,546 (7,276)	9,558 (7,288)	9,634 (7,360)	9,686 (7,388)
AR(1)	10,032 (8,276)	10,101 (8,440)	10,154 (8,561)	10,181 (8,613)	10,191 (8,627)	10,313 (8,822)	10,389 (8,943)	10,559 (9,182)
Até t-1								
LASSO	9,826 (7,286)	10,028 (7,677)	9,198 (7,286)	9,216 (7,263)	9,816 (7,834)	9,897 (7,852)	10,333 (8,896)	10,523 (8,913)
adaLASSO + OLS	11,661 (8,929)	10,267 (8,259)	9,726 (7,839)	9,746 (7,852)	10,303 (8,441)	10,222 (8,333)	9,634 (7,360)	9,686 (7,388)
AR(1)	10,032 (8,276)	10,101 (8,440)	10,154 (8,561)	10,181 (8,613)	10,191 (8,627)	10,313 (8,822)	10,389 (8,943)	10,559 (9,182)
Até t-2								
LASSO	8,207 (5,944)	8,460 (6,231)	8,724 (6,394)	9,026 (6,501)	8,380 (5,226)	8,379 (5,304)	8,207 (5,233)	8,629 (5,735)
adaLASSO + OLS	7,709 (4,927)	7,916 (5,259)	7,916 (5,067)	8,411 (5,660)	8,577 (5,749)	8,274 (5,744)	8,265 (5,657)	8,277 (6,235)
AR(1)	10,032 (8,276)	10,101 (8,440)	10,154 (8,561)	10,181 (8,613)	10,191 (8,627)	10,313 (8,822)	10,389 (8,943)	10,559 (9,182)
Até t-3								
LASSO	8,938 (6,900)	9,098 (7,225)	8,394 (6,548)	9,157 (7,168)	10,137 (8,092)	8,978 (6,808)	9,445 (7,333)	10,592 (8,612)
adaLASSO + OLS	8,496 (6,115)	9,589 (7,123)	8,517 (6,132)	8,918 (6,409)	10,054 (7,754)	10,045 (7,493)	10,074 (7,842)	9,738 (7,422)
AR(1)	10,032 (8,276)	10,101 (8,440)	10,154 (8,561)	10,181 (8,613)	10,191 (8,627)	10,313 (8,822)	10,389 (8,943)	10,559 (9,182)

(conclusão)

Início da Janela Amostral	2004/T1	2004/T2	2004/T3	2004/T4	2005/T1	2005/T2	2005/T3	2005/T4
Com variável <i>mobility</i>								
Sem lags								
LASSO	10,085 (8,636)	9,798 (8,308)	9,882 (8,261)	9,925 (8,317)	9,955 (8,345)	10,059 (8,492)	10,067 (8,386)	10,177 (8,506)
adaLASSO + OLS	10,010 (8,219)	9,332 (7,129)	9,600 (7,374)	9,531 (7,262)	9,546 (7,276)	9,558 (7,288)	9,634 (7,360)	9,686 (7,388)
AR(1)	10,032 (8,276)	10,101 (8,440)	10,154 (8,561)	10,181 (8,613)	10,191 (8,627)	10,313 (8,822)	10,389 (8,943)	10,559 (9,182)
Até t-1								
LASSO	11,550 (8,716)	10,028 (7,677)	9,198 (7,286)	9,216 (7,263)	9,816 (7,834)	9,897 (7,852)	10,333 (8,896)	10,523 (8,913)
adaLASSO + OLS	11,661 (8,929)	10,267 (8,259)	9,726 (7,839)	9,746 (7,852)	10,303 (8,441)	10,222 (8,333)	9,634 (7,360)	9,686 (7,388)
AR(1)	10,032 (8,276)	10,101 (8,440)	10,154 (8,561)	10,181 (8,613)	10,191 (8,627)	10,313 (8,822)	10,389 (8,943)	10,559 (9,182)
Até t-2								
LASSO	8,207 (5,944)	8,460 (6,231)	8,724 (6,394)	9,026 (6,501)	9,204 (6,188)	10,377 (6,844)	9,962 (6,622)	8,630 (5,739)
adaLASSO + OLS	7,709 (4,927)	7,916 (5,259)	7,916 (5,067)	8,411 (5,660)	8,577 (5,749)	8,274 (5,744)	8,265 (5,657)	8,277 (6,235)
AR(1)	10,032 (8,276)	10,101 (8,440)	10,154 (8,561)	10,181 (8,613)	10,191 (8,627)	10,313 (8,822)	10,389 (8,943)	10,559 (9,182)
Até t-3								
LASSO	8,938 (6,900)	9,098 (7,225)	8,394 (6,548)	9,157 (7,168)	10,137 (8,092)	8,977 (6,804)	9,445 (7,330)	10,592 (8,612)
adaLASSO + OLS	8,496 (6,115)	9,589 (7,123)	8,517 (6,132)	8,918 (6,409)	10,054 (7,754)	10,057 (7,563)	10,082 (7,887)	9,738 (7,422)
AR(1)	10,032 (8,276)	10,101 (8,440)	10,154 (8,561)	10,181 (8,613)	10,191 (8,627)	10,313 (8,822)	10,389 (8,943)	10,559 (9,182)

Notas: A tabela mostra a raiz do erro quadrático médio e o erro médio absoluto (entre parênteses) das previsões dos 64 conjuntos estimados. Os valores em negrito representam o melhor modelo de acordo com cada medida de erro e para cada data de início da janela amostral. Para cada data de início da janela amostral, permitimos a inclusão ou não da variável *mobility* entre as variáveis explicativas e a inclusão de até 3 defasagens