

**INSTITUTO TECNOLÓGICO DE AERONÁUTICA**



**Jaqueline Tortora Dias**

**ESTUDOS ECONOMÉTRICOS DOS IMPACTOS  
DO ICMS NO TRANSPORTE AÉREO E NA  
PRESERVAÇÃO DO MEIO AMBIENTE**

Trabalho de Graduação  
2020

**Curso de Engenharia Civil-Aeronáutica**

CDU: [ ]

**Jaqueline Tortora Dias**

**ESTUDOS ECONOMÉTRICOS DOS IMPACTOS  
DO ICMS NO TRANSPORTE AÉREO E NA  
PRESERVAÇÃO DO MEIO AMBIENTE**

Orientador

Prof. Dr. Alessandro Vinícius Marques de Oliveira (ITA)

**ENGENHARIA CIVIL-AERONÁUTICA**

SÃO JOSÉ DOS CAMPOS  
INSTITUTO TECNOLÓGICO DE AERONÁUTICA

2020

**Cataloging-in Publication Data**  
**Documentation and Information Division**

Dias, Jaqueline  
Estudos Econométricos dos Impactos do ICMS no Transporte Aéreo e na Preservação do Meio Ambiente / Jaqueline Dias.  
São José dos Campos, 2020.  
78p.

Trabalho de Graduação – Engenharia Civil-Aeronáutica – Instituto Tecnológico de Aeronáutica, 2020. Orientador: Prof. Dr. Alessandro Vinícius Marques de Oliveira

1. Impactos do ICMS na indústria de Aviação Civil e no Meio Ambiente. 2. Regressão LASSO. 3. Modelo de Preço. 4. Modelo de Demanda. 5. Aviação Civil. 6. Emissão de Gases Poluentes. I. Instituto Tecnológico de Aeronáutica. II. Estudos Econométricos dos Impactos do ICMS no Transporte Aéreo e na Preservação do Meio Ambiente.

## **REFERÊNCIA BIBLIOGRÁFICA**

DIAS, Jaqueline Tortora. **Estudos Econométricos dos Impactos do ICMS no Transporte Aéreo e na Preservação do Meio Ambiente**. 2020. 78p. Trabalho de Conclusão de Curso (Graduação em Engenharia Civil-Aeronáutica) – Instituto Tecnológico de Aeronáutica, São José dos Campos.

## **CESSÃO DE DIREITOS**

NOME DO AUTOR: Jaqueline Tortora Dias

TÍTULO DO TRABALHO: Estudos Econométricos dos Impactos do ICMS no Transporte Aéreo e na Preservação do Meio Ambiente

TIPO DO TRABALHO/ANO: Graduação / 2020

É concedido ao Instituto Tecnológico de Aeronáutica permissão para reproduzir cópias deste trabalho de graduação a para emprestar ou vender cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte deste trabalho de graduação pode ser reproduzida sem autorização do autor.

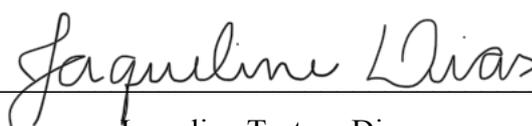
---

Jaqueline Tortora Dias  
R. H8A, 108 – Campus do CTA  
12228-460, São José dos Campos - SP



# ESTUDOS ECONOMÉTRICOS DOS IMPACTOS DO ICMS NO TRANSPORTE AÉREO E NA PRESERVAÇÃO DO MEIO AMBIENTE

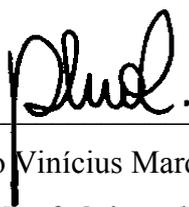
Essa publicação foi aceita como Relatório Final de Trabalho de Graduação



---

Jaqueline Tortora Dias

Autor



---

Prof. Dr. Alessandro Vinícius Marques de Oliveira (ITA)

Prof. Orientador



---

Prof. João Cláudio

Coordenador do Curso de Engenharia Civil-Aeronáutica

São José dos Campos, 18 de Novembro de 2020

A meus amigos e familiares que me ajudaram neste trabalho e em toda minha graduação.

## **Agradecimentos**

Ao ITA e todos os professores por todo o conhecimento que me ajudou a concluir esse projeto e a graduação.

Ao orientador deste trabalho Prof. Dr. Alessandro Vinícius Marques de Oliveira por todas as discussões e orientações durante todo este ano que foram fundamentais para o desenvolvimento deste trabalho.

*"Ó, vós que entraís, abandonai toda a esperança."*

Inferno (Divina Comédia)

Dante Alighieri

## Resumo

O presente trabalho propõe uma análise por econometria da correlação entre o ICMS sobre o querosene de aviação e os impactos que o aumento ou diminuição desse imposto podem gerar no setor de aviação e no meio ambiente. Em 2020 a humanidade enfrentou uma crise econômica e humanitária sem precedentes, uma pandemia que trouxe à tona discussões sobre a globalização e seus impactos sobre o meio ambiente e sobre a sociedade como um todo. Dentre os vários setores afetados, fez-se notável o grande impacto no setor aviação. Segundo dados da ANAC a demanda por voos no mercado doméstico teve queda de 85% em junho de 2020 em comparação com junho de 2019. Diante desse contexto o questionamento sobre como colaborar não só para a retomada do setor, mas para sua sobrevivência, se tornam frequentes e não esporadicamente envolvem o governo e as tributação sobre o setor. Devido à relevância do imposto sobre o combustível de aviação e o fato de ser pauta constante nas discussões das companhias e dos usuários, nesse trabalho vamos verificar a relação entre o ICMS e seus impactos no preço do combustível de aviação, e conseqüentemente, sobre o preço das passagens aéreas e as possíveis implicações dessa relação. O questionamento sobre as implicações do ICMS sobre o setor pode se converter em possíveis soluções para a atual situação. Porém, nesse mesmo contexto em que os hábitos da sociedade e seus impactos ao meio ambiente são constantemente colocados à prova, a pressão ambiental sobre a emissão de gases proveniente da operação de aviação civil é outro tema importante em pauta, e se a expectativa sobre o impacto do ICMS sobre o preço do combustível se mostrar verdadeira a dualidade que surge é o desenvolvimento econômico e recuperação da crise versus a preservação do meio ambiente. As conseqüências provenientes da posição do governo sobre a dualidade supracitada tem impactos de curto prazo sobre as estratégias das companhias do setor e sobre o desenvolvimento econômico nacional, tornando ainda mais proeminente o posicionamento do governo sobre o que é prioridade para a nação e seus representantes: a preservação e o futuro do planeta ou o desenvolvimento econômico.

## **Abstract**

This paper proposes an econometric analysis of the correlation between the ICMS on aviation fuel and the impacts that the increase or decrease of this tax can generate in the aviation sector and in the environment. In 2020, humanity faced an unprecedented economic and humanitarian crisis, a pandemic that brought up discussions about globalization and its impacts on the environment and society as a whole. Among the various sectors affected, the major impact on the aviation sector was notable. According to ANAC data, the demand for flights in the domestic market fell by 85% in June 2020 compared to June 2019. Given this context, questions about how to collaborate not only for the resumption of the sector but for its survival become frequent and they do not sporadically involve the government and taxation on the sector. Due to the relevance of the aviation fuel tax and the fact that it is a constant topic in the discussions of the companies and users, in this work we will verify the relationship between the ICMS and its impacts on the price of aviation fuel, and consequently, on the price airline tickets and the possible implications of this relationship. The questioning about the implications of the ICMS for the sector can become a solution to the current situation. However, in the same context in which society's habits and its impacts on the environment are constantly put to the test, the environmental pressure on the emission of gases from the civil aviation operation is another important issue on the agenda, and whether the expectation about the impact of the ICMS on the price of fuel proves to be true the duality that arises is the economic development and recovery from the crisis versus the preservation of the environment. The consequences arising from the government's position on the abovementioned duality have short-term impacts on the strategies of companies in the sector and on national economic development, even more prominent the government's position on what is a priority for the nation and its representatives to preserve and the future of the planet or economic development.

## Lista de Figuras

Figure 1: ICMS por estado e por companhia aérea no Brasil, 2019 - ABEAR.....	17
Figura 2 : Gráfico com histórico de preços do Golfo EUA para o combustível de aviação, gasolina, diesel, barril de petróleo (wti) e barril de petróleo (Brent) (legenda nessa ordem de cima para baixo) .....	19
Figura 3: Gráfico com o preço do barril de petróleo (Brent) (U\$/barril) e do preço do combustível de aviação (U\$/barril) .....	20
Figura 4: Preços dos combustíveis derivados do petróleo no Brasil, dados da IATA para 2012. ....	21
Figura 5: <i>Spread</i> no preço do combustível de aviação no Brasil .....	21
Figure 6: Itens que constituem o preço do combustível de aviação no Brasil. ....	22
Figura 7: Performance diária do preço da ação das principais companhias aéreas que operam no Brasil desde Janeiro de 2020. ....	24
Figura 8: Tradeoff entre viés e variância (Burnham & Anderson, 2004, p. 31).....	29
Figure 9: Descritivo da base de dados Stataer, 2020.....	31
Figure 10: Figura ilustrativa de um trecho do descritivo da base de dados Stataer, 2020. ....	38

## Lista de Tabelas

Tabela 1: Quadro de variáveis para o modelo do QAV	40
Tabela 2: Quadro de variáveis para o modelo de preço das passagens	42
Tabela 3: Quadro de variáveis para o modelo de demanda	45
Tabela 4: Regressão LASSO, nova seleção de variáveis e melhora nos resultados	49
Tabela 5: Regressão LASSO, ajuste de seleção de variáveis	50
Tabela 6: Regressão LASSO para Modelo de Demanda	52
Tabela 7: Inferência sobre possíveis efeitos do ICMS no preço do QAV e no preço das passagens	54

## Lista de Abreviações e Siglas

LASSO	Least Absolute Shrinkage and Selection Operator
QAV	Querosene de Aviação
ICMS	Sigla para o imposto sobre circulação de mercadorias e prestação de serviços de transporte interestadual e intermunicipal e comunicação
PIB	Produto Interno Bruto
ABEAR	Associação Brasileira das empresas Aéreas
IATA	Associação Internacional de Transporte Aéreo
ANAC	Agência Nacional de Aviação Civil
ANP	Agência Nacional de Petróleo
IBGE	Instituto Brasileiro de Geografia e Estatística
IPEA	Instituto de Pesquisa Econômica Aplicada
NECTAR	Núcleo de Economia do Transporte Aéreo
PLATTS	É um fornecedor de informações sobre energia e commodities e uma fonte de avaliações de preços de referência nos mercados físicos de commodities
OLS	Ordinary Least Squares
SINTAC	Sistema Integrado de Aviação Civil
IPi	Imposto sobre Produtos Industrializados
IRPJ	Imposto de Renda Pessoa Jurídica
CSLL	Contribuição Social sobre o Lucro Líquido
IRRF	Imposto de Renda Retido na Fonte
Cide	Contribuição de Intervenção no Domínio Econômico
IOF	Imposto sobre Operações Financeiras
IPVA	Imposto sobre a Propriedade de Veículos Automotores

ISS	Imposto Sobre Serviços
IPTU	Imposto Predial e Territorial Urbano
HHI	Índice Herfindahl-Hirschman

## Sumário

<b>INTRODUÇÃO</b> .....	<b>16</b>
<b>1.1 Motivação e Contexto</b> .....	<b>16</b>
<b>1.2 Definição do problema</b> .....	<b>19</b>
1.2.1 A dicotomia entre a recuperação do setor de aviação e o meio ambiente .....	22
1.2.2 Estrutura do trabalho .....	26
<b>2 REGRESSÃO LASSO</b> .....	<b>27</b>
<b>2.1 Introdução</b> .....	<b>27</b>
<b>2.2 O Processo de Regularização</b> .....	<b>29</b>
<b>2.3 Métodos de Encolhimento</b> .....	<b>30</b>
<b>2.4 Validação Cruzada (Cross-validation)</b> .....	<b>31</b>
<b>2.5 Efeitos Fixos</b> .....	<b>32</b>
2.5.1 Aplicações de efeitos fixos .....	32
2.5.2 Modelo de efeitos fixos .....	32
2.5.3 A importância da regressão de efeitos fixos .....	33
<b>2.6 Controles</b> .....	<b>34</b>
<b>3 METODOLOGIA</b> .....	<b>35</b>
<b>3.1 Construção dos modelos</b> .....	<b>35</b>
<b>3.2 Base de Dados</b> .....	<b>36</b>
<b>3.3 Modelo de Preço do Querosene de Aviação e o ICMS</b> .....	<b>38</b>
3.3.1 Variáveis Explicativas Adotadas no Modelo .....	38
<b>3.4 Modelo de Preço das Passagens Aéreas e o Querosene de Aviação</b> .....	<b>40</b>
3.4.1 Variáveis Explicativas Adotadas no Modelo .....	40
<b>3.5 Modelo de Demanda</b> .....	<b>42</b>
3.5.1 Variáveis Explicativas Adotadas no Modelo .....	42
<b>4 RESULTADOS</b> .....	<b>45</b>
<b>4.1 Modelo de Preço do Querosene de Aviação</b> .....	<b>46</b>
<b>4.2 Modelo de Preço da Passagem Aérea</b> .....	<b>49</b>
<b>4.3 Modelo de Demanda</b> .....	<b>50</b>
<b>4.4 Discussão de Resultados</b> .....	<b>53</b>
<b>5 CONCLUSÃO</b> .....	<b>55</b>
<b>6 BIBLIOGRAFIA</b> .....	<b>57</b>

# Introdução

## 1.1 Motivação e Contexto

As proporções da crise econômica global decorrente da pandemia do Covid-19 foram observadas no Brasil com a expectativa de queda de 5% no PIB nacional em 2020, divulgada em setembro de 2020 pelo Banco Central do Brasil, no Relatório Trimestral de Inflação. Essa estimativa é mais pessimista que a do Ministério da Economia, que prevê uma queda de 4,7% do PIB no mesmo período.

Além das consequências macroeconômicas de uma recessão, os impactos no setor de aviação civil foram ainda mais severos na atual crise: em uma pandemia em que o contato e a proximidade humana representam riscos letais à saúde, viagens e encontros presenciais cederam espaço à interação digital. Grande parte das companhias cessaram por tempo indeterminado algumas rotas e todas diminuíram as opções de voos pelo período de quarentena, forçando grandes empresas aéreas a realizar demissões em massa e até mesmo a entrar em recuperação judicial, como foi o caso da Latam Airlines.

O setor de aviação civil no Brasil gerou em 2019 R\$ 103,4 bilhões de valor adicionado a economia, equivalente a 1,4% do PIB nacional de acordo com a ABEAR, o volume de empregos atingiu 1,4 milhões, além de ser provedor de uma necessidade básica de locomoção entre longas distâncias, especialmente importante para um país de proporções continentais como o Brasil. Mitigar os efeitos adversos de uma crise sem precedentes e garantir a perenidade do setor é fundamental, dado o importante papel socioeconômico do segmento, fazendo com que um posicionamento do governo seja eminente e necessário.

Dentre as principais alternativas de estímulos no curto prazo destacam-se a oferta de crédito a taxas de juros subsidiadas e a implementação de incentivos fiscais. De acordo com informações do Senado Federal Brasileiro a aviação civil brasileira também sofre com a chamada tributação em cascata, imposto sobre imposto, nas três esferas de arrecadação – federal, estadual e municipal.

No âmbito federal, incidem o Imposto sobre Produtos Industrializados (IPI) na aquisição de novas aeronaves, mesmo as da Embraer; o Imposto de Renda Pessoa Jurídica (IRPJ) sobre o lucro real; a Contribuição Social sobre o Lucro Líquido (CSLL); o Imposto de Renda Retido na Fonte (IRRF) sobre o arrendamento internacional de aeronaves e peças (leasing) e as remessas para pagamento de serviços no exterior; a Contribuição de Intervenção no Domínio

Econômico (Cide) sobre os combustíveis; o Imposto de Importação sobre peças e componentes; e o Imposto sobre Operações Financeiras (IOF); entre outros previstos em lei. Os estados cobram o ICMS sobre insumos, arrendamento de aeronaves, transporte de carga e comércio de passagens aéreas, e o Imposto sobre a Propriedade de Veículos Automotores (IPVA) sobre a frota de automóveis das companhias aéreas. Finalmente, os municípios cobram o Imposto Sobre Serviços (ISS) na exportação de insumos e serviços pelas empresas aéreas; e o Imposto Predial e Territorial Urbano (IPTU) sobre lojas próprias.

Em 2001, a carga tributária comprometia 37% de toda a receita da aviação civil brasileira, ante 17% nos EUA e 9% na França, conforme estudo do Ipea de 2008, acrescentando que, pela estimativa para a carga tributária em relação ao PIB de 2009, é possível afirmar que a arrecadação no setor aéreo está acima da média dos setores produtivos.

Na média o querosene de aviação representa 30% dos custos das companhias aéreas brasileiras, no mundo esse valor está em torno de 20 a 22%, os valores foram publicados pela ABEAR e são referente ao ano de 2019. Na figura a seguir temos o mapa com os valores de ICMS cobrados em cada um dos estados brasileiros.

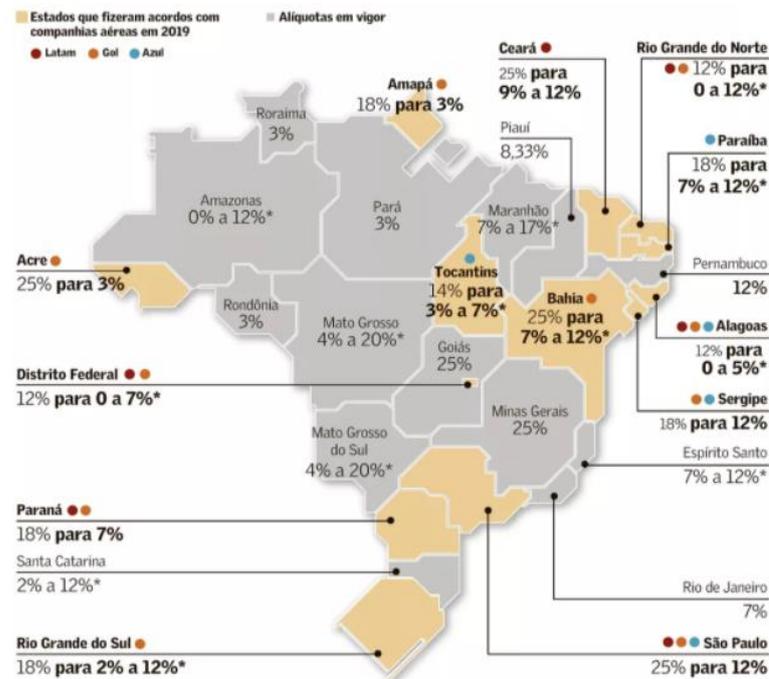


Figure 1: ICMS por estado e por companhia aérea no Brasil, 2019 - ABEAR

Importante ressaltar que, atualmente, o número de trabalhos que abordam a tributação na área de aviação são limitados em número, apesar da relevante contribuição do setor para o

PIB nacional e da significativa contribuição para emissão de gases prejudiciais ao meio ambiente. A contribuição deste trabalho seria fomentar a discussão por meio de teste de hipótese, a discussão já está em estágio mais avançado em países com a malha aérea mais desenvolvida, como Estados Unidos e os países europeus, como Reino Unido e Holanda. Pode-se notar pelas matérias citadas aqui previamente, o levantamento e comparação dos efeitos de tributos, como isso afeta a economia e meio ambiente de forma concomitante, o efeito da perda de emprego que o aumento de impostos ocasiona versus a diminuição de gases, análise comparativa citada no Financial Times, já referenciada neste trabalho.

O presente trabalho tem como principal objetivo testar as hipóteses de relação, possíveis impactos que uma decisão de redução do ICMS poderia gerar na aviação civil em termos econômicos e também os impactos no meio ambiente que essas alterações poderiam causar, para isso, estudam-se, de forma detalhada: (i) a relação entre o ICMS e o preço do combustível, paralelamente à relação entre o preço do combustível e o preço da passagem aérea, bem como à relação entre o preço da passagem, a demanda de passageiros e o volume de voos; (ii) outras possíveis relações entre preço, demanda, volume de passageiros e frequência de voos; (iii) uma vez estudada por modelo estatístico a expectativa de relação entre o ICMS e o preço da passagem aérea, discussão sobre potenciais impactos que a redução na alíquota do imposto estadual teria sobre a indústria da aviação civil e sobre políticas públicas de preservação do meio ambiente.

Para desenvolvimento de cada um dos questionamentos mencionados no parágrafo anterior, realiza-se o procedimento a seguir: contextualização e discussão do problema real, modelagem econométrica do problema, explicação do modelo teórico utilizado, análise dos resultados obtidos, refinamento e alterações no modelo proposto após a observação dos primeiros resultados e, por fim, conclusão com base nos resultados obtidos do modelo refinado.

Na primeira seção, definem-se os principais objetos de análise: os impactos da carga tributária de ICMS no preço do querosene de aviação e as consequências na indústria de aviação civil e no meio ambiente.

Na seção subsequente, que trata das ferramentas econométricas utilizadas, detalha-se o modelo LASSO, as variáveis causais e investigativas, os efeitos fixos, controles e os métodos de encolhimento. Após apresentação de visão inicial da utilização de LASSO, analisam-se as primeiras relações entre as variáveis causais definidas, para após validação dessas relações construir o modelo com as variáveis investigativas e então introduzir efeitos fixos e controles para refinamento do modelo por meio da análise de resultados.

## 1.2 Definição do problema

Vamos detalhar a precificação do querosene de aviação, objeto de estudo, para esclarecimento e contextualização do assunto. Retomando a importância da indústria de aviação civil para economia global. A indústria foi estimada em 2019 em U\$ 838 bilhões<sup>1</sup>, conectando 4,2 bilhões de pessoas em 2018 segundo a mesma fonte. Os principais benefícios de uma indústria de aviação civil bem desenvolvida são o aumento do PIB gerado, os empregos, e os impostos recolhidos dessa indústria, desconsiderando outros impactos positivos diretos e indiretos da indústria. Apesar da grande importância, as margens das companhias aéreas e da indústria de aviação civil como um todo são relativamente baixas, com custos altos e tributação relevante no custo total das companhias aéreas.

Um custo representativo no custo total das companhias aéreas é o combustível de aviação, que é produzido por destilação fracionada do petróleo. No processo de produção a gasolina de aviação vai competir com gasolina e diesel, o que contribui para o aumento do preço. Na figura a seguir temos o gráfico em que essa relação pode ser observada.

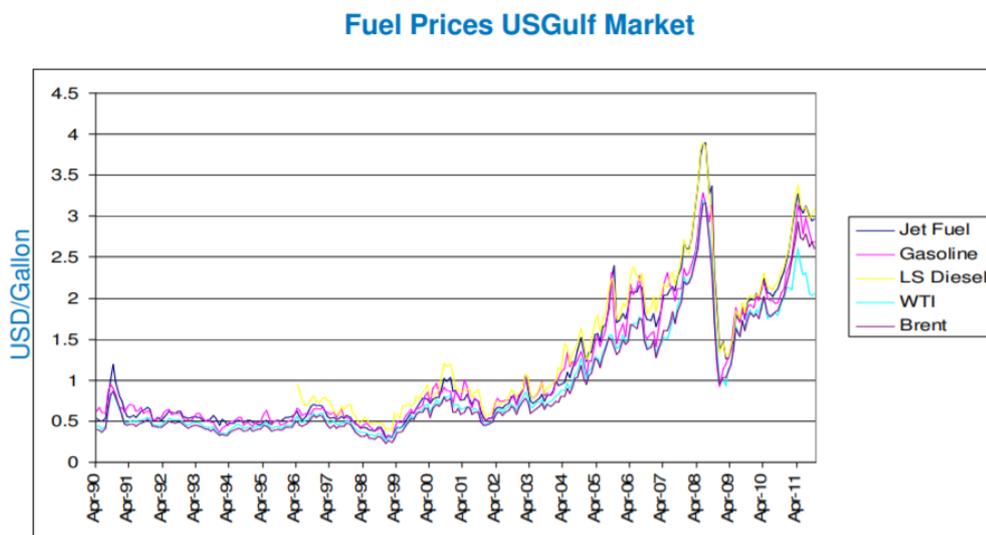


Figura 2 : Gráfico com histórico de preços do Golfo EUA para o combustível de aviação, gasolina, diesel, barril de petróleo (wti) e barril de petróleo (Brent) (legenda nessa ordem de cima para baixo)

<sup>1</sup> Dados do Banco Mundial e IATA em Outubro de 2020.

Para a precificação do querosene de aviação, a maioria dos países utiliza um preço base para o preço do querosene de aviação. Existem três principais companhias que fornecem esse preço base: Platts, OPIS e Argus. Essas companhias coletam dados públicos de transações de combustível de aviação e publicam uma média diariamente. Dessas companhias os valores mais utilizados com referência são os da Platts, cerca de 90% do mercado global. A Platts coleta e publica os dados em 11 regiões do mundo, Golfo (EUA), Los Angeles (EUA), São Francisco (EUA), Nova Iorque (EUA), Chicago (EUA), Rotterdam, MEDAV (Itália), Europa Ocidental, Oriente Médio, Cingapura e Japão. Para a América Latina, Caribe e Sul dos Estados Unidos a publicação utilizada como preço base é a Platts para o Golfo EUA.

Os contratos entre as distribuidoras e as companhias aéreas incluem cláusulas específicas com o preço público base que será utilizado para cada local ou aeroporto. Os preços divulgados pelo Platts já incluem os custos do óleo cru, refinaria, transporte, armazenagem e margem de lucro, ou seja, se a distribuidora utilizar o preço base ela já estará obtendo lucro. Veja a seguir a Figura 2 com o histórico do preço do barril de petróleo e do combustível de aviação.

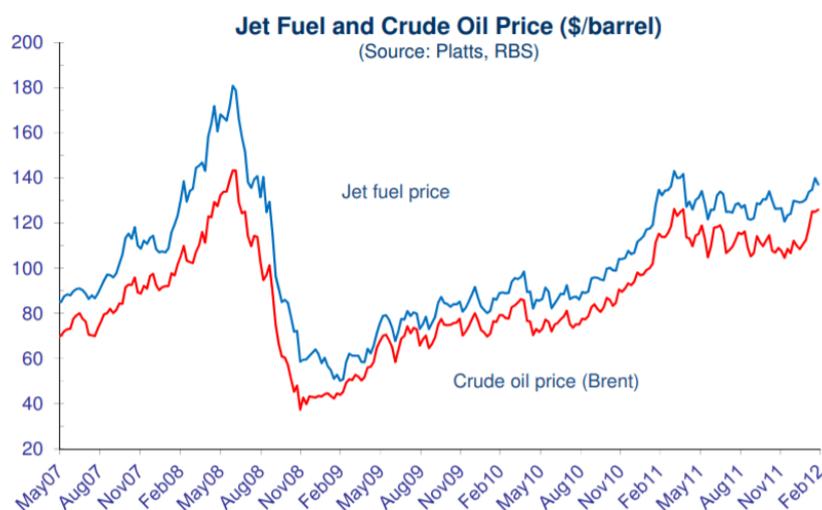


Figura 3: Gráfico com o preço do barril de petróleo (Brent) (US\$/barril) e do preço do combustível de aviação (US\$/barril)

De posse dessas informações iniciais, vamos analisar o contexto brasileiro de precificação do querosene de aviação (QAV). Importante ressaltar que o combustível é o custo operacional mais relevante para as companhias aéreas. Em uma média mundial esse custo

representa cerca de 34%<sup>2</sup> dos custos operacionais das companhias aéreas. No Brasil esse custo está em torno de 40% dos custos operacionais para as companhias aéreas que operam no país. Os preços do combustível de aviação no Brasil são, em média, 12% maiores que na América Latina e 17% maiores que a média global. Veja nas figuras 4 e 5 a seguir os preços de combustíveis, em geral no Brasil e o *spread*, ou seja, a cobrança adicional no preço definido pelo Platts que ocorre no Brasil.

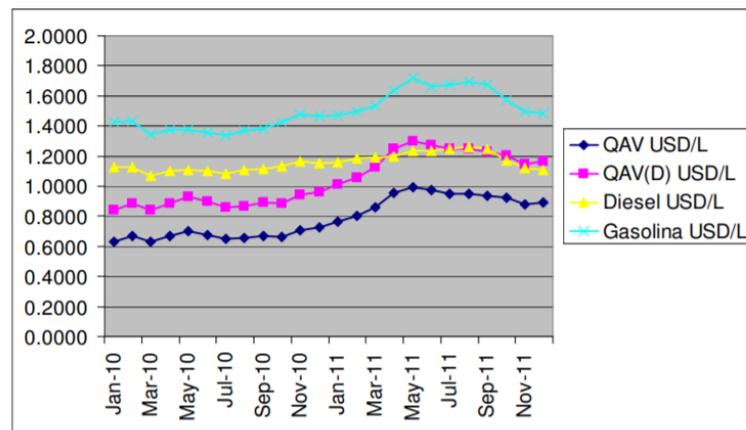


Figura 4: Preços dos combustíveis derivados do petróleo no Brasil, dados da IATA para 2012.

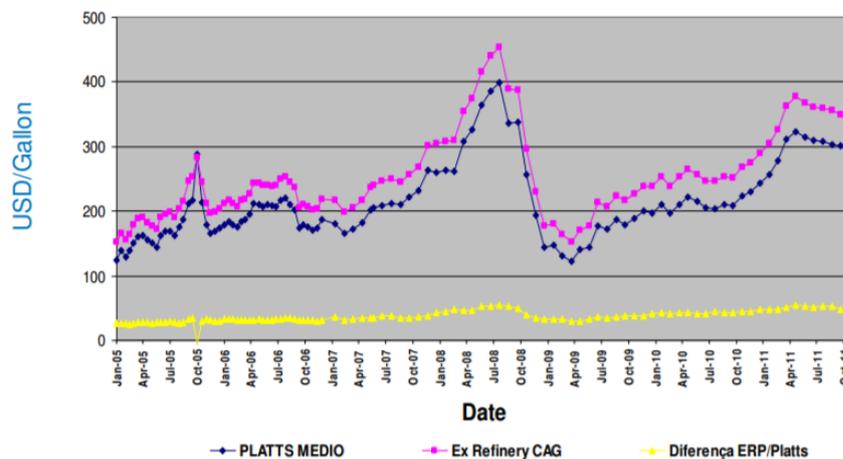


Figura 5: *Spread* no preço do combustível de aviação no Brasil

Finalmente, a construção do preço do combustível no Brasil no gráfico da Figura 6 a seguir, que é composto basicamente pelo preço definido pelo Platts e o impostos, como pode ser observado pelas curvas azul (Platts médio) e roxa (Preço com ICMS).

<sup>2</sup> Dados da IATA em 2014.

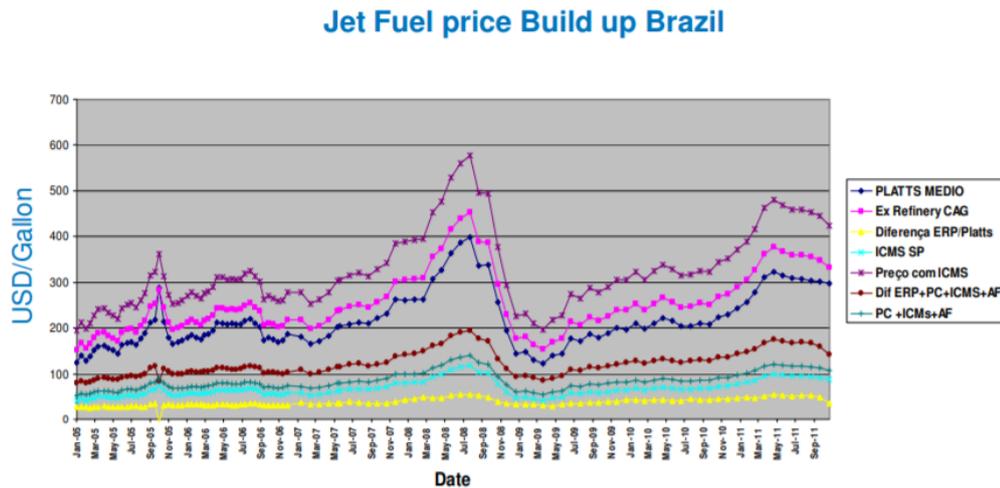


Figure 6: Itens que constituem o preço do combustível de aviação no Brasil.

Ao verificar por meio de modelos teóricos aplicados a dados reais a significância da relação entre o preço das passagens e o ICMS pode-se formular políticas públicas que influenciem diretamente no setor, seja para impulsioná-lo, seja para criar mecanismo de proteção ao meio ambiente por meio de aumento de impostos.

Nesse contexto a pesquisa será restrita ao ICMS, apesar de outros impostos também influenciarem nos resultados da análise descrita no parágrafo anterior. Dessa forma a aplicação do método LASSO foi escolhida para, justamente, isolar a análise e dirimir efeitos adversos de outras variáveis. Vale ressaltar que o LASSO é uma metodologia de *machine learning*, o machine learning é um conjunto de técnicas que vem fortalecendo a previsão com aplicação na econometria, pois no caso do LASSO, por exemplo, a metodologia é capaz de aperfeiçoar o modelo por meio da análise de dados de experiência, a própria metodologia LASSO torna o modelo mais simples, diminuindo o viés sem perder nível de variância adequado.

### 1.2.1 A dicotomia entre a recuperação do setor de aviação e o meio ambiente

O objetivo principal deste trabalho é verificar as relações do ICMS com o preço do combustível, o preço das passagens aéreas e uma possível relação que esses resultados podem ter com a demanda. Ao testar a tendência da literatura quanto aos impactos das medidas de redução desse impostos sobre a indústria de aviação pode-se delimitar possíveis medidas de colaboração entre governo e o setor e, mais importante, relacionar essa retomada com o meio ambiente e com todos os desdobramentos que a alteração do imposto pode ter, das diferentes perspectivas que essa alteração pode ser analisada.

Medidas restritivas adotadas em todas as regiões do globo afetaram fortemente o setor de aviação. O relatório semestral divulgado pela IATA reporta que “o ano de 2020 será o pior ano da história para as companhias aéreas”. Segundo o relatório, as perdas devem ficar em torno de US\$ 84,3 bilhões. No Brasil, o impacto da Covid-19 na economia nacional será forte. Um estudo de conjuntura econômica do IPEA publicado no dia 9 de Outubro prevê uma redução de 6% do PIB para o ano de 2020. As empresas aéreas como Gol e Azul registram no 1º trimestre desse ano uma deterioração de patrimônio líquido de R\$ 3,3 bilhões e R\$ 6,13 bilhões, respectivamente.

Provavelmente, as perdas do setor de aviação devem ser maiores no segundo trimestre. No mês de abril foram suspensos os voos internacionais e o mercado presenciou uma queda de 93% na demanda por voos domésticos. Com as restrições de mobilidade sendo prolongadas em todas as regiões do país e do mundo, qualquer expectativa de recuperação do setor será adiada para o segundo semestre de 2020. Ademais, os baixos índices de confiança direcionam para uma recuperação lenta e gradual das atividades econômicas.

A pandemia gerou impactos severos no setor de aviação brasileiro, comparando o mês de abril de 2020 com o mesmo mês de 2019 observa-se para o setor de aviação: (i) Redução de 90,9% no número de decolagens; (ii) Diminuição de 94,5% no número de passageiros; (iii) Queda de 16,3% na taxa de ocupação das aeronaves; (iv) As ações da GOL (GOLL4) sofreram queda de 45,6% e da Azul (AZUL4) queda de 48,8%. Na Figura 7 é possível observar a enorme queda no valor de mercado dessas companhias comparando ao início da quarentena, com uma retomada lenta e ainda muito distante dos patamares de antes da pandemia.

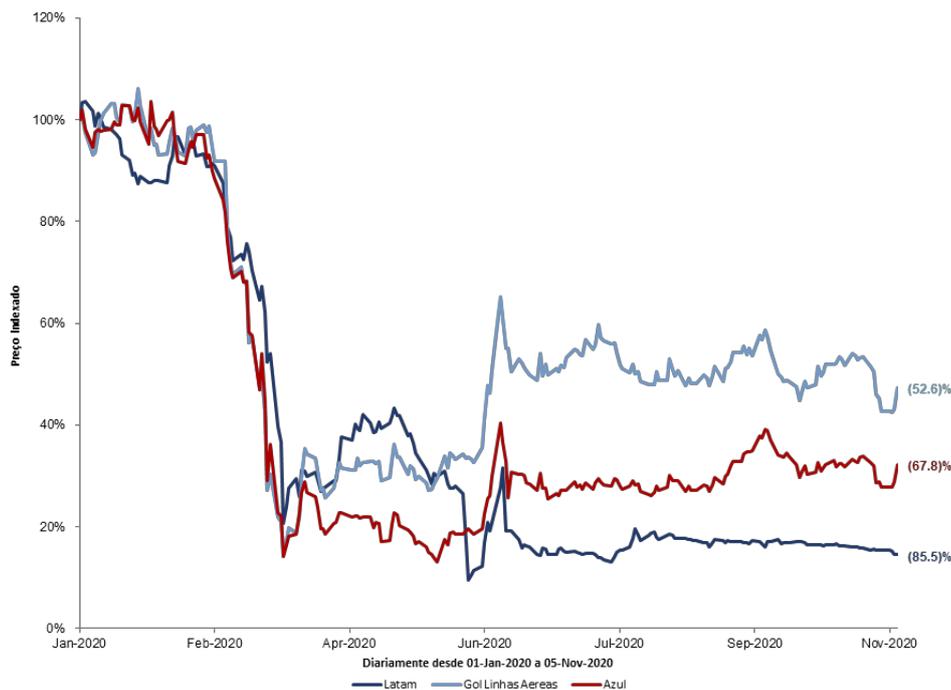


Figura 7: Performance diária do preço da ação das principais companhias aéreas que operam no Brasil desde Janeiro de 2020.

Uma medida de ajuda por parte do governo que pode ser adotada é o corte ou redução do ICMS sobre o querosene de aviação por determinado período de tempo. As questões que podem surgir é se a medida diminui os custos das companhias aéreas de maneira significativa de modo que a redução na arrecadação seja capaz de colaborar com a recuperação das companhias aéreas e do setor.

Outra questão importante sobre um possível corte remete a uma tendência mais forte na Europa relacionada à taxa de passageiros e companhias aéreas devido a grande emissão de gases poluentes provenientes da indústria de aviação. De acordo com matéria veiculada no jornal *Financial Times* de Maio de 2019 o setor gera cerca de 2,5% das emissões globais de dióxido de carbono, de acordo com um documento apresentado em fevereiro pelas autoridades holandesas à União Européia. As emissões aumentam com o crescimento econômico e devem aumentar em 68% entre 2010 e 2020. A Holanda organizou uma conferência internacional em 20 de junho de 2019 para discutir preços de carbono e impostos de aviação, abrindo espaço para a discussão da abrangência de taxa sobre todos os setores, para que a contribuição seja efetiva e a luta contra as mudanças climáticas seja pauta central nas discussões econômicas do bloco. A Holanda também faz parte da Coalizão de Ministros de Finanças para Ação Climática, formada em abril de 2019 para ajudar a financiar iniciativas nacionais de ação climática,

compartilhar ideias e incluir riscos climáticos em planos econômicos. O grupo tem 26 membros, incluindo 11 estados membros da UE - Áustria, Dinamarca, Finlândia, França, Alemanha, Irlanda, Luxemburgo, Holanda, Espanha, Suécia e Reino Unido.

Na Europa, o combustível de aviação custa US\$ 85 o barril, tendo estado a maior parte de 2018 entre os valores de US\$ 80 - 90 e quase chegando a US\$ 100 no início de outubro de 2018. Valores bem maiores que os que prevaleciam três anos atrás, cerca de US\$ 50. A Platts, empresa fornecedora dos dados sobre o preço padrão do combustível, estimou no final do ano de 2017 que o alto preço do combustível custaria às companhias aéreas US\$ 50 bilhões adicionais em 2018.

Em discussões mais recentes, a dicotomia levanta as duas opções, necessidade de cortar impostos para ajudar na recuperação e a oportunidade de uma retomada limpa, esta segunda tem ganhado força entre os especialistas do setor. Em matéria publicada no *New York Times* em março de 2020, foi apontada a queda nas viagens devido ao vírus, que também pode prejudicar um antigo plano das Nações Unidas de limitar as emissões de voos internacionais aos níveis de 2020, já que os níveis de 2020 serão muito abaixo da média de anos com atividade normal das companhias aéreas. Esse plano exigiria que as companhias aéreas compensassem o aumento das emissões nos anos futuros, financiando projetos que reduzam a emissão do dióxido de carbono ou o colaborem para sua retirada da atmosfera, por exemplo, plantando árvores ou instalando energia renovável.

Mas o problema que surge é, se 2020 for um ano excepcionalmente baixo para as emissões das companhias aéreas - como é muito provável que seja, por causa do medo do coronavírus - isso poderia aumentar significativamente os custos para as companhias aéreas nos anos futuros se as viagens voltarem aos níveis tradicionais. Uma análise preliminar do Fundo de Defesa Ambiental descobriu que, no segmento de ponta, as companhias aéreas podem precisar comprar até 20% mais créditos do que esperavam nos próximos três anos, dependendo de quanto as viagens aéreas diminuíam e se recuperem.

Grupos ambientalistas disseram que o surto não deve ser usado como uma desculpa para enfraquecer programas projetados para lidar com as emissões da aviação. “Este é um momento desafiador para as companhias aéreas, mas isso não é motivo para enfraquecer essas regras e desviar as companhias aéreas dos esforços para lidar com a mudança climática”, disse Annie Petsonk, conselheira internacional do Fundo de Defesa Ambiental na mesma entrevista ao *Financial Times* citada previamente.

A aviação é responsável por 2,4% das emissões mundiais de dióxido de carbono (CO<sub>2</sub>), e a demanda por voos deve dobrar nas próximas 2 décadas, comparativamente aos níveis normais

de 2019. Se as atuais metas de descarbonização da indústria não forem cumpridas, a aviação pode contribuir com até um quarto de todo o orçamento de carbono do mundo até 2050. Se o imposto tem papel relevante no assunto é plausível assumir que atrelar o imposto às medidas de contenção de emissão de gás carbônico, ou mesmo exigir restituição das companhias aéreas desse imposto por meio de iniciativas desenvolvidas dentro da própria companhia (como plantar árvores ou investir em políticas públicas de locomoção não poluente) é indispensável e deve ocorrer o quanto antes.

### **1.2.2 Estrutura do trabalho**

Este trabalho será detalhado ao decorrer dos capítulos subsequentes, retomando as principais ideias abordadas neste capítulo. Além disso os métodos utilizados serão desenvolvidos nos capítulos 3 e 4, e conclusões serão obtidas no capítulo 5.

Uma revisão do método LASSO é realizada no capítulo 2, por se tratar de uma ferramenta de modelo econométrico ainda pouco utilizada na academia brasileira, e pelo caráter teórico de grande importância para a pesquisa em questão. O capítulo 3 contempla a análise da base de dados e restrições ao modelo para obtenção de resultados contudentes. Na sequência, a metodologia e modelos utilizados são seguidos pela apresentação de resultados, com discussão sobre os possíveis esclarecimentos que as relações obtidas podem gerar e como esses dados podem ser avaliados comparativamente ao problema real.

A obtenção de resultados e análises sobre as consequências desses resultados são realizadas no capítulo 4, que aliado aos capítulos teóricos 2 e 3 permite a elaboração de uma conclusão no capítulo 5.

## 2 Regressão Lasso

### 2.1 Introdução

Econometria é um dos mais populares métodos quantitativos para explicação de fenômenos humanos e comportamentos da sociedade. Sendo universalmente conhecido e utilizado de maneira simplista ou, muitas vezes, inapropriada. Com o avanço da tecnologia e aumento expressivo da quantidade de dados que é possível coletar sobre um fenômeno, as técnicas de estudo foram evoluindo e atualmente uma regressão linear simples não é mais suficiente para chegar a conclusões dada a complexidade do sistema e dos problemas estudados. Porém vamos partir do básico e avançar para modelos mais sofisticados para esclarecer os avanços e o aumento de confiabilidade nos resultados obtidos pela metodologia que utilizaremos neste trabalho.

Ao se estudar um fenômeno qualquer, a econometria tem como principais objetivos o entendimento e previsão. Esse entendimento passa por um processo de estimação de relações entre as variáveis e como estão relacionadas dentro da análise, de forma a modelar a situação entender os mecanismos que podem desencadear alguma alteração no problema que se está modelando. Muitas vezes o melhor ajuste não é suficiente e nem mesmo o objetivo final, por exemplo, a equação que descreve perfeitamente uma determinada amostra de dados pode não explicar muito bem uma nova amostra de um mesmo fenômeno e isso demonstra o não entendimento do problema em questão.

A ideia geral do modelo clássico é decompor o fenômeno em seus fatores determinantes. Quais fatores são determinantes para o fenômeno que queremos entender ou descrever é o estudo que vamos desenvolver neste trabalho. Não se sabe previamente todos os fatores que são capazes de modificar um determinado fenômeno, o que se faz é estudá-lo para determinar o maior número de características que se encaixa na maior parte das situações em que o fenômeno ocorre, para que mesmo com a atualização ou adição de dados para o fenômeno o entendimento seja amplo o suficiente para nos fornecer informações importantes sobre o fenômeno mas que também seja capaz de se adequar a diferentes situações do mesmo.

Existem várias escolas de pensamento na metodologia econométrica, apresentamos aqui o tradicional ou metodologia clássica, que ainda domina a pesquisa empírica em economia e

outras ciências sociais e comportamentais<sup>3</sup>. Em termos gerais, a metodologia econométrica tradicional prossegue ao longo das seguintes linhas:

- I. Declaração de teoria ou hipótese.
- II. Especificação do modelo matemático da teoria
- III. Especificação do modelo estatístico ou econométrico
- IV. Obtenção dos dados
- V. Estimativa dos parâmetros do modelo econométrico
- VI. Teste de hipóteses
- VII. Previsão ou predição
- VIII. Usando o modelo para fins de controle ou política.

Desse modelo clássico dois desafios principais surgem ao tentar se construir modelos com base de dados com alto número de observações. Observa-se que o aumento do número de variáveis na equação de um modelo aumenta também o  $R^2$ , sem necessariamente aumentar a confiabilidade do modelo, o que torna o acréscimo de variáveis algo aparentemente benéfico para o modelo, porém o que ocorre é o *overfitting* da curva de dados e, conseqüente, o modelo decora a base de dados e não funcionará da forma correta ao se inserir novos dados, o chamado viés.

Viés e variância constituem os principais desafios na regressão linear clássica. De tal forma que um modelo com número elevado de variáveis irá se adaptar perfeitamente à curva, porém, quando ocorre a adição de novos dados e alteração da base o modelo já não explica muito bem devido ao excesso de variáveis que foram adicionadas à equação para se aproximar ao máximo da base de dados inicial, o chamado *overfitting*. Da mesma forma o pesquisador pode diminuir de forma significativa o número de variáveis, então a equação será muito diferente da curva de dados e explicará poucas ou nenhuma das características da base de dados, o chamado *underfitting*. Metodologias mais complexas como Ridge e Lasso, os chamados métodos de encolhimento buscam um ponto ótimo entre os dois desafios listados com foco em evitar o *overfitting*, como mostra a Figura 8:

---

<sup>3</sup> Damodar N. Gujarati, Basic Econometrics, © The McGraw-Hill Companies, 2004, p.3.

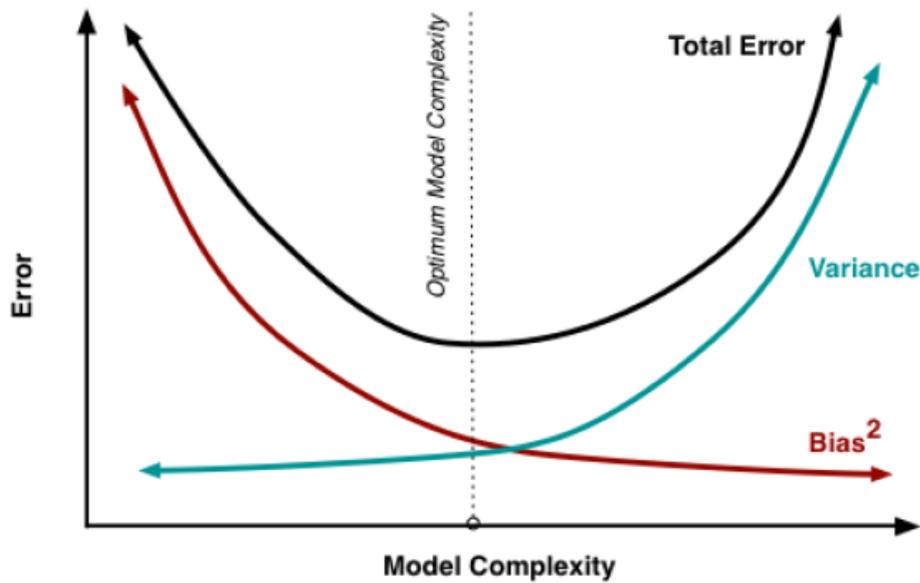


Figura 8: Tradeoff entre viés e variância (Burnham & Anderson, 2004, p. 31).

## 2.2 O Processo de Regularização

A regularização é um procedimento de otimização que impõe restrições à estimação dos Mínimos Quadrados Ordinários (Ordinary Least Squares. OLS), de tal forma que incentiva uma solução caracterizada pela atenuação dos efeitos da complexidade inerente aos modelos de alta dimensão (HD) – por meio ou não de esparsidade<sup>4</sup>. No caso de aplicar soluções esparsas, o faz pela introdução de maior parcimônia na especificação, prevenindo o sobreajuste dos dados. Também conhecido como “método de regressão penalizada”, sendo um tipo de Mínimos Quadrados Restritos.

A esparsidade por sua vez é caracterizada como assumir previamente que a maioria das variáveis não é significativa para o modelo, principalmente em modelos HD. A simplificação adotada no conceito de esparsidade parte do pressuposto de duas ideias básicas: (i) a explicação mais simples de um problema é a mais provável de estar correta e (ii) não conseguimos explicar de maneira coerente problemas com muitas variáveis, não sendo muito útil um modelo com muitas variáveis de qualquer forma.

<sup>4</sup> Esparsidade: para uma base de dados com número elevado de variáveis pode-se dizer que esparsidade se refere ao fenômeno de que “uma estrutura de dados subjacente pode ser explicada principalmente por poucos de muitos recursos” – Sirimongkolkasem & Drikvandi (2019).

## 2.3 Métodos de Encolhimento

Para Ridge e Lasso o que ocorre é o encolhimento de variáveis que, por meio de checagem automática, variáveis que não são essenciais para o modelo são eliminadas ou tem sua dimensão minimizada, são principalmente aplicáveis para reduzir o número de variáveis, selecionar as que são realmente essenciais, de forma a simplificar o modelo. Para isso adiciona-se um termo à equação principal de regressão linear, essas equações podem ser observadas a seguir:

$$\text{Ridge: } \hat{\beta}^{ridge} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^n (y_i - \beta_0 - \sum_j \beta_j x_{ij})^2 \right\}, \text{ sujeita à } \sum_j \beta_j^2 \leq t \quad (1)$$

$$\text{Lasso: } (\hat{\alpha}, \hat{\beta}) = \underset{\alpha, \beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^n (y_i - \beta_0 - \sum_j \beta_j x_{ij})^2 \right\}, \text{ sujeita à } \sum_j |\beta_j| \leq t \quad (2)$$

Tibshirami, 1996 - Regression Shrinkage and Selection via the Lasso

De tal forma que a quantidade do encolhimento é controlada pelo parametro  $t \geq 0$  que é aplicado nas estimativas. A diferença mais importante entre Lasso e Ridge é que o Lasso é capaz de eliminar variáveis, enquanto Ridge apenas as encolhe. Neste trabalho vamos utilizar o Lasso, pois este método funciona também como aperfeiçoador do modelo por meio de seleção de variáveis.

LASSO é a sigla para “Least Absolute Shrinkage and Selection Operator”, de tradução, “Operador de Encolhimento Absoluto Mínimo e Seleção”. A metodologia do LASSO que proporciona ao usuário da técnica obter os parâmetros de mínimos quadrados, técnica usualmente utilizada, também realiza de forma simultânea o encolhimento de coeficientes, ou seja, a redução do valor do coeficiente refletindo a significância da variável no modelo. Além de reduzir coeficientes o LASSO também é capaz de zerá-los, realizando dessa forma uma seleção dos regressores caracterizando-o como um bom método de inferência estatística, pois o LASSO seleciona o melhor modelo por meio de análise das variáveis e não por simples escolha do usuário.

Na metodologia LASSO vamos ter o regressando (Variável dependente, Variável de “desfecho”), os regressores ou variáveis explicativas que serão constituídos por dois grupos essenciais as variáveis causais, inquestionáveis no modelo devido a sua relação já ter sido comprovada na literatura (consenso entre os estudiosos), e as variáveis explicativas, serão as

variáveis que explicarão de fato as relações que se quer verificar, além dessas temos os efeitos fixos e os controles que serão explicados mais detalhadamente em tópicos futuros neste mesmo capítulo.

## 2.4 Validação Cruzada (Cross-validation)

A validação cruzada tem como objetivo otimizar o desempenho de previsão fora da amostra do modelo, é uma técnica para avaliar modelos por meio de treinamento de vários subconjuntos de dados de entrada disponíveis e avaliação deles no subconjunto complementar dos dados. Usa-se a validação cruzada para detectar sobreajuste, ou seja, a não generalização de um padrão.

O método mais conhecido de execução da validação cruzada é o k-fold. Na validação cruzada k-fold, você divide os dados de entrada em subconjuntos de dados k (também chamados de *fold*s, pastas). Você treina um modelo em todos esses subconjuntos, exceto um deles (k-1) e depois avalia o modelo nesse subconjunto que não foi usado para o treinamento. Esse processo é repetido k vezes, com um subconjunto diferente reservado para avaliação (e excluído do treinamento) a cada vez.

Veja a seguir o diagrama que mostra um exemplo de subconjuntos de treinamento e subconjuntos de avaliação complementar gerados para cada um dos quatro modelos que são criados e treinados durante uma validação cruzada 4-fold. O modelo 1 usa os primeiros 25% dos dados para avaliação e os 75% restantes para treinamento. O modelo dois usa o segundo subconjunto de 25 por cento (25 a 50 por cento) para avaliação, e os três subconjuntos restantes de dados para treinamento e assim por diante.

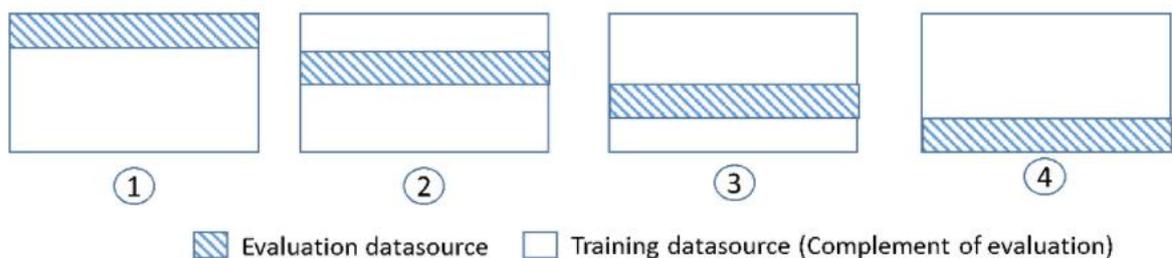


Figure 9: Descritivo da base de dados Stataer, 2020.

## 2.5 Efeitos Fixos

### 2.5.1 Aplicações de efeitos fixos

Um modelo de efeitos fixos se refere a um modelo de regressão no qual as médias do grupo são fixas (não aleatórias) em oposição a um modelo de efeitos aleatórios em que as médias do grupo são uma amostra aleatória de uma população. Geralmente, os dados podem ser agrupados de acordo com vários fatores observados. As médias do grupo podem ser modeladas como efeitos fixos ou aleatórios para cada agrupamento. Em um modelo de efeitos fixos, cada média do grupo é uma quantidade fixa específica do grupo.

Esses modelos ajudam a controlar o viés de variável omitida devido à heterogeneidade não observada quando esta heterogeneidade é constante ao longo do tempo. Essa heterogeneidade pode ser removida dos dados por meio da diferenciação, por exemplo, subtraindo a média do nível de grupo ao longo do tempo ou tomando uma primeira diferença que removerá quaisquer componentes invariáveis de tempo do modelo.

Existem duas suposições comuns feitas sobre o efeito específico individual: a suposição de efeitos aleatórios e a suposição de efeitos fixos. A suposição de efeitos aleatórios é que os efeitos específicos do indivíduo não estão correlacionados com as variáveis independentes. A suposição de efeitos fixos é que os efeitos específicos do indivíduo estão correlacionados com as variáveis independentes. Se a suposição de efeitos aleatórios for mantida, o estimador de efeitos aleatórios é mais eficiente do que o estimador de efeitos fixos. No entanto, se essa suposição não for válida, o estimador de efeitos aleatórios não é consistente. Vale ressaltar que a hipótese dos efeitos aleatórios tende a ser irrealista, vamos trabalhar com efeitos fixos na construção dos resultados neste trabalho.

### 2.5.2 Modelo de efeitos fixos

Considere o modelo linear de efeitos não-observáveis para N observações e T períodos de tempo:

$$y_{it} = X_{it}\beta + \alpha_i + u_{it} \text{ for } t = 1, \dots, T \text{ e } i = 1, \dots, N \quad (3)$$

Em que:

$y_{it}$  é a variável dependente observada para o indivíduo  $i$  no tempo  $t$

$X_{it}$  é o variante no tempo  $1 \times k$  (número de variáveis independentes)

$\beta$  é a  $1 \times k$  matriz de parâmetros

$\alpha_i$  é o efeito individual não-observável e invariante no tempo

$u_{it}$  é o termo de erro

De tal forma que  $X_{it}$  e  $\alpha_i$  não podem ser observados diretamente, já que no modelo de efeitos fixos  $\alpha_i$  pode estar correlacionado com a matriz regressor  $X_{it}$ .

Temos então o estimador de efeitos fixos, como  $\alpha_i$  não pode ser observado diretamente, também não pode ser diretamente controlado, o modelo de efeitos fixos elimina o  $\alpha_i$  degradando as variáveis por meio de transformação interna.

$$y_{it} - \bar{y}_i = (X_{it} - \bar{X}_i)\beta + (\alpha_i - \bar{\alpha}_i) + (u_{it} - \bar{u}_i) \Rightarrow \dot{y}_{it} = \dot{X}_{it}\beta + \dot{u}_{it} \quad (4)$$

Onde  $\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}$ ,  $\bar{X}_i = \frac{1}{T} \sum_{t=1}^T X_{it}$  e  $\bar{u}_i = \frac{1}{T} \sum_{t=1}^T u_{it}$ , como  $\alpha_i$  é constante e  $\bar{\alpha}_i = \alpha_i$  e dessa forma o efeito é eliminado. O  $\beta$  do estimador de efeitos fixos é então obtido por uma regressão OLS de  $\dot{y}$  e  $\dot{X}$ .

Neste trabalho a transformação interna considera o valor médio da característica dos grupos como a característica de todos os indivíduos. A fim de controlar os fatores específicos individuais que sejam invariantes no tempo, estima-se os efeitos fixos de acordo com o procedimento especificado na equação (4), em geral, não os sujeitando à penalização no procedimento LASSO.

### 2.5.3 A importância da regressão de efeitos fixos

As regressões de efeitos fixos são muito importantes porque os dados muitas vezes se enquadram em categorias como indústrias, estados, regiões, famílias, etc. Normalmente deseja-se controlar as características dessas categorias que podem afetar o regressando. Não há como ter certeza de que se possui todas as variáveis de controle relevantes, então, se você estimar um modelo OLS simples, terá que se preocupar com fatores não observáveis que estão correlacionados com as variáveis que você incluiu na regressão.

Se esses fatores não observáveis são invariantes no tempo, então os efeitos fixos eliminarão o viés da variável omitida. Em alguns casos, pode-se acreditar que seu conjunto de variáveis de controle é suficientemente rico para quaisquer partes não observáveis do ruído de regressão e, portanto, o viés da variável omitida se aproxima do zero. Mas como já mencionado

não há como ter certeza sobre isso, já que esse fatores são inobserváveis, portanto, os modelos de efeitos fixos são uma boa precaução.

## 2.6 Controles

Também denominada “variáveis de incômodo “ (*nuisance variables*) são variáveis que devem ser estatisticamente controladas no estudo para promover a inferência de causalidade das variáveis de interesse. São caracterizadas como fatores que interferem na análise de outro e, em geral, não possuem relevância direta para o problema. O controle colabora para garantir o efeito *ceteris paribus* na análise, ou seja, os efeitos isolados das variáveis de interesse.

A estimação de parâmetros associados à diversas variáveis, cada uma com suas idiosincrasias, pode não ser interessante para o modelo pois afeta o efeito isolado de outras variáveis que são as de principal interesse no estudo - por isso são chamados de “*nuisance parameters*”. Essas variáveis servem apenas com um recurso que aproxima o modelo de uma função desconhecida, não-paramétrica, e assim controlar efeitos não observáveis para viabilizar uma identificação consistente do efeito da variável de interesse. Quando são variáveis de alta dimensão, algum procedimento de redução da dimensionalidade deve ser aplicado aos controles utilizados, sendo o LASSO um modelo de Pós-Regularização/Seleção para se “procurar fatores de confusão” nos dados.

## 3 Metodologia

### 3.1 Construção dos modelos

A modelagem econométrica segue um determinado padrão para garantir sua eficácia na inferência de ideias a partir da análise de dados. Seguindo os padrões de econometria determina-se o fenômeno a ser estudado, como o objetivo deste trabalho é entender o impacto do ICMS na indústria de aviação, este é o objeto de estudo e ao formular-se perguntas de pesquisa sobre o assunto, ou seja, dúvidas reais que alguém que necessita entendimentno sobre o tema, percebe-se a grande necessidade de simplificação, portanto faz-se um recorte do fenômeno e das perguntas relacionadas a ele para conseguir inferências sobre o assunto.

Definido o fenômeno, perguntas sobre as questões reais que envolvem o fenômeno são elaboradas para construção do modelo conceitual que irá explicar, para uma versão simplificada do fenômeno, essas perguntas. Utiliza-se funções de variáveis, em que as variáveis representarão um conceito teórico da realidade. Partindo de um conhecimento prévio, de estudos já elaborados e que já são consenso entre os estudiosos do assunto, ou seja, um conhecimento já estabelecido, constroem-se as hipóteses que serão validadas com alicerces nesse conhecimento estabelecido.

A base de dados utilizada será detalhada em tópico posterior, e representa a amostra da realidade onde o modelo será testado e após refinamento, validado. A base requer muita observação das diferentes fontes e de sua confiabilidade, conhecimento (experiência no setor, termos e jargões) e levantamentos (histórico de fatos, legislação, políticas, mudanças em discussão, órgãos reguladores), sendo item fundamental para qualidade da inferência.

O fenômeno a ser estudo neste trabalho são as relações entre o querosene de aviação, o ICMS, o preço das passagens aéreas, a demanda por passagens aéreas e como isso afeta a indústria aeronáutica, e o modelo conceitual envolve todas essas variáveis.

Para construção dos modelos de análises utilizados neste estudo foram consideradas etapas para análises prévias ao Lasso e, conseqüentemente, comparativas. Dessa forma a primeira abordagem foi rodar uma regressão simples considerando os efeitos fixos. Após a primeira abordagem a regressão foi considerada com efeitos fixos, para então serem escolhidas e determinadas as dummies. A segunda abordagem inclui essas dummies juntamente com os efeitos fixos já considerados. Na Terceira abordagem, juntamente às dummies e aos efeitos fixos insere-se os controles de presença. Num quarto momento, para construção do Lasso,

calcula-se a validação cruzada para num quinto momento, de posse de todos os elementos necessários para utilização do Lasso calcula-se o Lasso com a consideração dos efeitos fixos e penalização dos controles e por fim, no sexto momento, cálculo do Lasso penalizando controles e variáveis exógenas.

As abordagens de construção da tabela final comparativa de sensibilidade foi utilizada nos três modelos construídos neste trabalho (i) o modelo de preço do querosene e ICMS; (ii) o modelo de preço da passagem aérea, o querosene de aviação e o ICMS; (iii) modelo de demanda e oferta incluindo também o ICMS e o querosene de aviação.

### **3.2 Base de Dados**

A base de dados utilizada para o desenvolvimento do presente trabalho está armazenada no *software* Stata e é uma iniciativa do Núcleo de Economia do Transporte Aéreo (NECTAR) do Instituto Tecnológico de Aeronáutica. Os valores são um compilado de dados fundamentais para construção de análises do setor de aviação brasileiro e provém de diferentes fontes oficiais e órgãos reguladores, destaca-se a Agência Nacional de Aviação Civil (ANAC), Agência Nacional de Petróleo (ANP), Instituto Brasileiro de Geografia e Estatística (IBGE) e o Instituto de Pesquisa e Economia aplicada (IPEA). A base ainda está em desenvolvimento e passa por constantes atualizações.

Importante ressaltar que os dados estão organizados por rota, par de cidades, definidas como origem e destino, e apresentam as informações para cada rota como, por exemplo: número de voos, distância, número de passagens vendidas na rota, o tipo de aeronave, entre outras informações.

No desenvolvimento deste projeto a descrição da base analisada continha 646 variáveis, sendo 96 defasadas – variáveis que foram atualizadas ou retiradas da base – restando 550 variáveis classificadas em 19 grupos listados a seguir: *\_classificadores*, *aircraft*, *costs*, *disruptions*, *distance*, *dummies*, *flights*, *market structure*, *national share*, *network*, *output*, *pax*, *pax connections*, *presence*, *price & revenue*, *shares & proportions*, *slots*, *socioeconomic*, *survey*. A descrição da base oferece 18 itens de reconhecimento da variável, como a classificação citada anteriormente, a fonte primária (órgão oficial onde foi coletada a informação) e breve descritivo de limitação da variável, item que fornece além da descrição as datas referentes aos dados. Verifica-se a distribuição de variáveis em cada uma dessas classes na Tabela 1.

Tabela 1: Classificação dos grupos de variáveis da Base NECTAR

<b>Grupos</b>	<b>Grupos - Português</b>	<b>Número de Variáveis</b>
_classificadores	classificadores	25
aircraft	aeronave	60
costs	custos	13
disruptions	disrupções	38
distance	distância	9
dummies	fictícios	13
flights	voos	28
market structure	estrutura de mercado	13
national share	participação nacional	15
network	rede	19
output	dados de saída	30
pax	passageiros	16
pax connections	conexões de passageiros	18
presence	presença	163
price & revenue	preço e receita	3
shares & proportions	participação & proporção	27
slots	vagas	13
socioeconomic	socioeconômico	25
survey	pesquisa	22
<b>Total</b>		<b>550</b>

Ao todo base conta com 206.433 observações e 1.176 variáveis, sendo 550 variáveis originais e as demais variações destas, como média geométrica, valores máximo e mínimos, entre outros.

A base é instrumento essencial para desenvolvimento do projeto e acuridade dos resultados, já que sua procedência e organização é essencial para um rápido reconhecimento e confiabilidade dos resultados obtidos por meio dos modelos econométricos construídos. A seguir a demonstração do descritivo da base Nectar com algumas das respectivas classificações fornecidas. Vale ressaltar que a base está em constante evolução e que novas versões estão sendo implementadas no Stata durante o desenvolvimento deste trabalho.

nome da variável	classe da variável	tipo de variável	fonte primária	cômputo da variável	level	descritivo 1	descritivo 2
_or_gdp_d	socioeconomic	contínua	IBGE	NECTAR	city	gross domestic product	produto interno bruto
_or_gdp_o	socioeconomic	contínua	IBGE	NECTAR	city	gross domestic product	produto interno bruto

Figure 10: Figura ilustrativa de um trecho do descritivo da base de dados Stataer, 2020.

Após estudo das 550 variáveis contidas na base, fez-se a seleção dos grupos de variáveis de interesse conforme classificação oferecida pelo descritivo da base. Para os modelos de preço e demanda que serão desenvolvidos neste trabalho as classificações que se mostraram mais relevantes foram: aircraft, costs, market structure, national shares, price & revenue, shares and proportion e socioeconomics. Partindo dessa seleção inicial de variáveis, construiu-se o refinamento dos modelos e seus desdobramentos.

De posse do estudo detalhado da base utilizada neste trabalho pode-se desenvolver os próximos passos do ferramental teórico utilizado nos modelos de análise.

### 3.3 Modelo de Preço do Querosene de Aviação e o ICMS

O primeiro modelo para estimativa da influência do ICMS no preço do querosene de aviação foi construído sobre a premissa básica de que o preço internacional do barril de petróleo convertido para reais é uma variável causal. Especificamente, utilizou-se as variáveis preço do dólar ‘usd\_o’ e o preço do barril de petróleo ‘wti\_o’ ambos com a classificação de origem para possibilitar análise comparativa dos dados e garantir uma relação coerente com o que se observa na realidade e também na literatura existente sobre o tema.

#### 3.3.1 Variáveis Explicativas Adotadas no Modelo

Para explicar os possíveis efeitos do ICMS sobre o preço do querosene de aviação foram selecionadas também outras variáveis para que sejam considerados de forma individual seus efeitos sobre o regressor.

- a) Média diária mensal em US\$ do preço do petróleo (wti) (neste modelo Média diária mensal do preço do petróleo em US\$ convertido para reais): a relação do preço do petróleo com o preço do querosene de aviação é utilizada neste modelo principalmente como medida de verificação do modelo, já que a relação já foi exemplificada neste trabalho e é um consenso na literatura. Vale mencionar que o preço do petróleo em dólares foi convertido para seu valor em reais usando a taxa de conversão de moedas disponível na base de dados utilizada.
- b) Etapa de voo (avstl): Média das etapas de voo do aeroporto de destino em km. Proxy para custos das companhias aéreas, com variabilidade temporal (depende do número de cidades servidas, que varia). Distância em km, geodésica Vincenty. A variável 'astl' foi incluída pela consideração de que quanto mais longa a etapa, ou, quanto mais afastado o local de destino do voo, mais caros tendem a ser os insumos, principalmente para cidade muito afastadas das capitais e grandes centros urbanos devido ao alto custo de deslocamento desses insumos ao local em questão.
- c) Igual a km vezes número de voos. Distância em km, geodésica Vincenty (flownkm): o entendimento da influência da distância no preço do querosene de aviação está relacionado com o conceito de disponibilidade e desenvolvimento da região em que o combustível será obtido. As regiões mais distantes costumam ser menos desenvolvidas e além disso é considerável que o preço dos combustível nessas regiões seja maior devido a custo de transporte. Adiciona-se a esse custo a acessibilidade e desenvolvimento da região considerada, que em geral costuma ser menos desenvolvida com menos consumo de combustível e portanto poucas oportunidades de negociação de preço comparativamente a regiões metropolitanas.
- d) Total de litros consumidos de combustível nos voos da rota, conforme obtido do SINTAC (fueltt): variável para avaliar a relação entre média de desempenho das aeronaves na rota e o consumo de combustível, e conseqüentemente, a relação com o preço do querosene de aviação.

- e) Tempo médio planejado de voo (em horas) (schfltime): escolha de variável por estar relacionada com o conceito de distância de destino, assim como explicado para variável ‘flownkm’.
- f) ICMS (gmicms): média geométrica entre o ICMS da origem e o ICMS do destino. Variável cuja relação com o regressando é o objetivo de pesquisa deste trabalho.

Na Tabela 2, apresenta-se um resumo das variáveis utilizadas no modelo de preço do QAV. Todas as considerações realizadas sobre as variáveis, as relações entre elas e como elas se comportam no modelo são expectativas que serão verificadas na seção de resultados.

Tabela 1: Quadro de variáveis para o modelo do QAV

<b>Variável</b>	<b>Descrição da Variável</b>
jal	Preço do querosene de aviação excluído o valor do ICMS
wti	Preço do barril de petróleo em dólar, convertido para reais
icms	Alíquota do ICMS sobre o QAV
flownkm	km percorridos em média por voo na rota vezes o número total de voos nesta rota
fuellt	Consumo médio de combustível dos voos da rota (em litros)
avstl	Distância média percorrida em cada etapa de voo (km)
schfltime	Tempo programado de voo

### **3.4 Modelo de Preço das Passagens Aéreas e o Querosene de Aviação**

No segundo modelo que estima a influência do ICMS no preço das passagens aéreas, considerou-se com variáveis causais o preço do querosene de aviação e o número total de voos, coerente com o que se observa na realidade e também com a literatura existente sobre o tema.

#### **3.4.1 Variáveis Explicativas Adotadas no Modelo**

Para identificar os possíveis efeitos do ICMS sobre o preço das passagens aéreas foram selecionadas também outras variáveis para que sejam considerados de forma individual seus

efeitos sobre o regressor. Especificamente, utilizou-se o preço médio das passagens aéreas<sup>5</sup> na rota, deflacionado de acordo com o IPCA sendo a base o mês de janeiro de 2019.

- a) Preço do QAV disponibilizado pela ANP, por estado/dia (fja1): Calculado como o menor preço entre a origem e o destino. Poderia ser o preço da origem apenas, a correlação é alta. Incorpora valor dos impostos. Deflacionado IPCA jan 2019. Novamente, a relação entre o preço do querosene de aviação e o preço da passagem aérea é um link entre o nosso primeiro modelo, em que 'ja1' era o regressando e a variável regressando da preço da passagem (p), além de ser variável causal de relação esperada positiva.
- b) Total de voos da rota (f): Importante salientar que o número total de voos é indicativo de aumento ou diminuição de preços por estar relacionado com o número de possibilidades de compra e com o número de possibilidades de rota, que adiciona liberdade de escolha ao usuário e pode influenciar, de forma indireta no preço. Atender um espectro maior de clientes e, possivelmente, aumentar o número de passageiros afetando também o preço.
- c) Índice de Herfindahl-Hirschman (gmhhi): O Índice Herfindahl-Hirschman (HHI) é uma medida comum de concentração de mercado e é usado para determinar a competitividade do mercado, muitas vezes antes e depois de transações de fusões ou aquisições entre companhias com grande parcela do mercado. O Índice Herfindahl-Hirschman (HHI) é uma medida comumente aceita de concentração de mercado. É calculado elevando-se ao quadrado a participação de mercado de cada empresa que compete em um mercado e, em seguida, somando os números resultantes. Pode variar de quase zero a 10.000. O Departamento de Justiça dos Estados Unidos (U.S. Department of Justice) usa o HHI para avaliar possíveis problemas de fusão. Além disso estudiosos do setor de aviação aérea utilizam o HHI considerando as 50 maiores empresas do setor para determinar se ele pode ser considerado competitivo ou mais próximo de ser um monopólio. O prefixo gm que precede a sigla hhi no modelo, e para

---

<sup>5</sup> A base de tarifas da ANAC somente acomoda valores detalhados a partir de julho de 2010. Períodos anteriores foram desconsiderados.; A base inclui todas as passagens aéreas vendidas no período, para viagens com qualquer antecedência.

todas as variáveis da base NECTAR, significa média geométrica entre origem e destino (*geometric mean* do inglês);

- d) Média de litros de combustível utilizados por quilômetro voado nos voos da rota, conforme obtido do SINTAC (fuelpkm): os litros de combustível consumidos foram admitidos no modelo para considerar os possíveis efeitos de uma frota menos eficiente e para verificar se o aumento de consumo em determinadas rotas influencia o preço do querosene de aviação;
- e) ICMS (gmicms): média geométrica entre o ICMS da origem e o ICMS do destino. Variável cuja relação com o regressante é o objetivo de pesquisa deste trabalho.

Na Tabela 3, apresenta-se um resumo das variáveis utilizadas no modelo de preço das passagens aéreas. Todas as considerações realizadas sobre as variáveis, as relações entre elas e como elas se comportam no modelo são expectativas que serão verificadas na seção de resultados.

Tabela 2: Quadro de variáveis para o modelo de preço das passagens

Variável	Descrição da Variável
p	Preço médio das passagens aéreas na rota
fjal	Preço do querosene de aviação somado ao valor do ICMS
f	Total de voos da rota
gmicms	Média geométrica entre ICMS de origem e destino
gmhhi	Herfindahl-Hirschman Índice, índice de concentração de mercado
fuelpkm	Média de litros consumidos de combustível por km voado

### 3.5 Modelo de Demanda

Para o terceiro modelo, temos uma análise da demanda. As variáveis causais consideradas foram o preço da passagem aérea e a média geométrica do PIB per capita para possibilitar análise comparativa dos dados e garantir uma relação coerente com o que se observa na realidade e também na literatura existente sobre o tema.

#### 3.5.1 Variáveis Explicativas Adotadas no Modelo

Para tentar explicar e construir a equação de demanda foram selecionadas também outras variáveis para que as variáveis causais e as variáveis explicativas de forma conjunta sejam capazes de nos fornecer um maior entendimento da demanda, para o regressando desse modelo foi escolhida a variável ‘tickets’.

- a) Número de passagens vendidas (tickets): a variável tickets fornece o número total de passagens vendidas<sup>6</sup>. É o número correspondente às passagens aéreas vendidas no período, para viagens com qualquer antecedência.
- b) Total de passageiros pagos transportados em voos regulares (rpax): variável essencial na estimativa de demanda, já que denota a demanda em termos de volume de passageiros transportados.
- c) Preço médio das passagens aéreas na rota (p): O valor é deflacionado pelo IPCA sendo a base janeiro de 2019. É esperado que a relação entre preço e demanda seja negativa, ou seja, que o aumento de um aumento de preço leve a uma redução na quantidade demandada. Este é um fator importante na análise de modelos de demanda.
- d) Total de voos da rota<sup>7</sup> (f): Importante salientar que o número total de voos é indicativo de aumento de demanda, é esperado uma correlação positiva, já que o aumento de oferta de voos em uma rota oferece mais possibilidades aos usuários, em termos de horários e conexões, ou seja, há uma tendência a atender a um espectro maior de clientes e, possivelmente, aumentar a demanda.
- e) Preço do QAV obtido da ANP, por estado/dia (ja1): Calculado como o preço mínimo da origem e destino. Poderia ser o preço da origem apenas, a correlação é alta. Não incorpora impostos. O valor é deflacionado pelo IPCA sendo a base janeiro de 2019. Novamente, a relação entre o preço do querosene de aviação e a demanda por passagens aéreas é um link entre o nosso primeiro modelo, em que ‘ja1’ era o regressando e o terceiro modelo, em que ‘ja1’ é a uma variável regressora.

---

<sup>6</sup> A base de tarifas da ANAC somente acomoda valores detalhados a partir de julho de 2010. Períodos anteriores foram desconsiderados.

<sup>7</sup> Matriz OD não abrange os passageiros que trocam de aeronave. Dados da etapa básica SINTAC. Com ajustes de outliers de km, asize e f (conciliação do cálculo).

- f) Proxy para PIB per capita no período; vide respectivas variáveis proxies de população e PIB para o período (gmgpc): o PIB per capita é o indicador macroeconômico da equação. Trata-se de um importante regressor, já que o poder de compra tem correlação positiva esperada alta. É esperado que quanto maior a renda maior o acesso ao transporte aéreo, já que é o de menor tempo e, geralmente, o menos acessível em termos de custos para o usuário.
- g) Número de cidades servidas (ncs): o desenvolvimento da malha aérea é um fator importante na determinação de demanda, já que quando mais lugares são atendidos pela malha aérea, maior o número de pessoas conectadas à malha aérea, e mais pessoas vão utilizar o meio de transporte.
- h) Passageiros em conexão (maxconn): os passageiros em conexão é medida que pode afetar a relação entre as demais variáveis, já que essa variável pode dar falsa impressão de demanda a depender do aeroporto e da malha aérea local.
- i) Assentos disponíveis voos fretados (charter) (sseats\_chart): é uma variável relacionada a demanda turística, inserida no estudo para verificar se a disponibilidade de assentos pode ser fator relevante para a demanda, pois entende-se que a variável pode influenciar de maneira significativa o preço.

Na Tabela 4, apresenta-se um resumo das variáveis utilizadas no modelo de demanda. Todas as considerações realizadas sobre as variáveis, as relações entre elas e como elas se comportam no modelo são expectativas que serão verificadas na seção de resultados.

Tabela 3: Quadro de variáveis para o modelo de demanda

Variável	Descrição da Variável
tickets	Número de passagens vendidas
rpax	Total de passageiros pagos transportados em voos regulares
p	Preço médio das passagens aéreas na rota
f	Total de voos da rota
jal	preço do querosene de aviação excluído o valor do ICMS
gmgpc	PIB per capita
maxconn	Número de passageiros em conexão
ncs	Número de cidades servidas
sseats_chart	Assentos disponíveis em voos fretados

## 4 Resultados

Os processos de construção dos 3 modelos que serão detalhados ao longo da seção 4 seguiram o diagrama da figura 11.

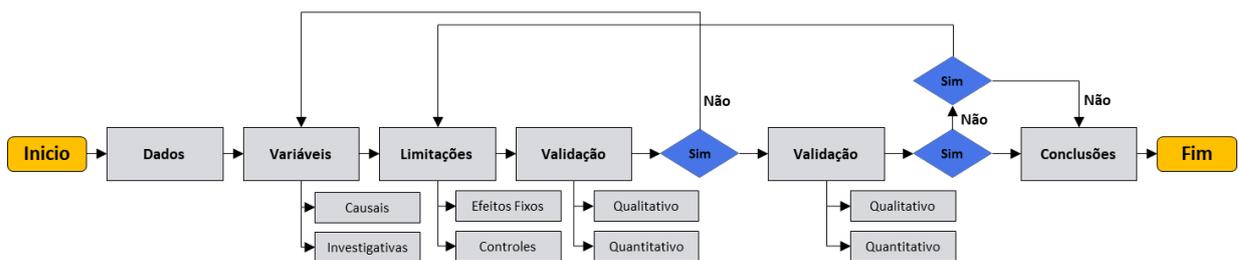


Figura 11: Diagrama de decisão para construção dos 3 modelos<sup>8</sup>

No início, preocupou-se em estudar os dados disponíveis na base de dados do Stater, para entender como foram organizados os dados, as classificações dos agrupamentos de variáveis e suas variações. Realizou-se avaliação inicial de como esses grupos se relacionam com o regressor de cada um dos modelos propostos e como se relacionam entre si.

Após esse estudo inicial dos dados fez-se seleção dos grupos de variáveis de interesse. Dentro desses grupos eliminou-se as que tinham relação colinear com o regressando e selecionou-se quais possuíam uma expectativa de relação com o regressando que fosse consenso na literatura, essas variáveis que aparentavam possuir relação de consenso foram definidas e introduzidas ao código como variáveis causais, todas as demais se tornaram variáveis que seriam introduzidas no modelo, a princípio, apenas para observação de seu comportamento. Todas essas variáveis, causais e investigativas, somaram 52 e foram

<sup>8</sup> Diagram inspirado em figura apresentada por Rafael Prado em apresentação parcial do trabalho de graduação **Ensaio Econométrico sobre eficiência energética de aeronaves no transporte aéreo regular brasileiro**, 2020.

selecionadas, para depois se introduzir os efeitos fixos e os controles com variáveis que possivelmente poderiam ocasionar ruído nos coeficientes dos regressandos.

Após a conclusão dos três passos detalhados: (i) dados, (ii) variáveis, e (iii) limitações, todas essas 52 variáveis foram introduzidas nos respectivos modelos para verificações iniciais de seus comportamentos, possíveis eliminações do LASSO e, principalmente, a validação das variáveis causais. Nesse momento houve a seleção manual de variáveis por meio dos resultados obtidos e refinamento da seleção em caso de eliminação do LASSO, não significância estatística, ou relação não condizente com o observado na realidade, consenso na literatura.

Esse processo se seguiu de maneira repetitiva, redefinindo a seleção de variáveis, avaliando resultados, se necessário, redefinindo as limitações e especificações de efeitos fixos e controles, verificando os resultados, até chegar a um modelo que demonstrasse de maneira satisfatória a relação entre regressores e regressandos.

Os resultados desses processos é a seleção do modelo que melhor explica essa relação entre regressor e regressando para cada um dos modelos propostos, esse modelos são explicados e seus resultados verificados no detalhamento de cada um deles nas subseções seguintes.

Vale ressaltar que o LASSO acelera esse processo de seleção das variáveis e refinamento do modelo, porém, o trabalho manual de checagem dos resultados e verificação de cada uma das variáveis é manual e demanda quantidade de tempo considerável.

Outro fato importante para as análises é que todas as variáveis estão em logaritmos neperianos, e por conta disso, suas dimensões são comparáveis.

#### **4.1 Modelo de Preço do Querosene de Aviação**

Como detalhado anteriormente no capítulo 3 deste trabalho, esse modelo tem como objetivo a análise do preço do querosene de aviação e sua relação com o ICMS, tendo como variáveis causais o preço do barril de petróleo e o dólar. Nos primeiros modelos de preço desse trabalho, analisou-se a relação entre o preço do querosene de aviação (variável regressando, variável dependente ou variável de “desfecho”) e o preço internacional do barril de petróleo em dólar (regressores ou variáveis explicativas) e a relação do mesmo regressor, preço do querosene de aviação, com o dólar.

Inicialmente o estudo foi realizado considerando separadamente essas variáveis, preço do barril de petróleo em dólar (variável ‘wti’ ou ‘wti\_o’ em que o sufixo ‘o’ indica que essa variável tem seus valores relacionados à origem da rota, ou par de cidades) e o dólar (variável

‘usd’ ou ‘usd\_o’ em que o sufixo ‘o’ indica que essa variável tem seus valores relacionados à origem da rota, ou par de cidades), porém essa análise não rendeu resultados satisfatórios e, dessa forma, optou-se por considerar o preço do barril de petróleo em reais. A conversão foi realizada com os valores de dólar da mesma base.

A hipótese inicial era de que a relação do preço do querosene de aviação e o preço do barril de petróleo é positiva e estatisticamente significativa, enquanto a relação com o dólar é negativa e estatisticamente significativa. De posse dessa informação, mantém-se a expectativa com relação ao preço do barril de petróleo em reais.

A variável ‘fope’ (representa o número total de voos operados, não cancelados), entre outras, foi incluída no modelo em versões iniciais, porém foi eliminada pelo LASSO, sendo retirada do modelo. A variável ‘rpax\_st’ (representa o total de passageiros pagos transportados em voos regulares) também foi considerada em versões iniciais, com resultados de não significância estatística e possuindo coeficiente de módulo muito menor que as demais variáveis, dessa forma também foi desconsiderada. Adicionalmente, nos modelos iniciais, o LASSO eliminou a variável ‘fuelpkm’ (média de litros consumidos de combustível por quilômetro voado nos voos da rota, conforme obtido do SINTAC). Em substituição ao ‘fuelpkm’ (média de litros consumidos de combustível por quilômetro voado nos voos da rota) utilizou-se a variável ‘flownkm’ (representa distância da rota em km vezes número de voos. Distância em km, geodésica Vincenty), que demonstrou maior consistência na análise de sensibilidade.

Nesse mesmo modelo, incluiu-se a variável investigativa de maior interesse nesse estudo, o ICMS (‘icms\_o’, em que o sufixo ‘o’ indica que os valores estão relacionados com a origem da rota, par de cidades, considerados), alterando as considerações iniciais com a variável ‘jal’ (representa o preço do combustível sem inclusão da alíquota do ICMS) para a variável de preço cheio do combustível de aviação (‘fja1’), ou seja, que inclui o valor do imposto.

A inferência esperada sobre o ICMS é de uma relação positiva entre o preço do querosene de aviação e o ICMS, ou seja, o aumento do ICMS provoca aumento no preço do combustível, assim como a diminuição do ICMS provoca queda no preço do querosene de aviação.

Sobre os resultados obtidos da variável ‘hhi\_o’ (representa índice de concentração do setor de aviação no Brasil, o sufixo ‘o’ significa que os dados são relativos a origem da rota, par de cidades), vale lembrar que o Brasil tem um mercado concentrado de distribuidores de QAV, portanto a competitividade entre as companhias aéreas e a própria isenção de impostos

que cada uma delas possui em cada estado pode ser fator determinante do preço. Pelos resultados obtidos dessas análises para definição do modelo final percebe-se relação positiva entre o ‘hhi\_o’ (representa índice de concentração do setor de aviação no Brasil) e o preço do QAV (‘ja1\_o’) com significância estatística a 1% e coeficientes de modulo relevante comparativamente aos demais. Porém, uma relação positiva entre preço e competitividade não se aplica ao que se observa na realidade, pois isso significa que o decréscimo dessa variável (resultado de um aumento de concorrência) aumenta o preço, ou seja, na verdade é esperada uma relação negativa, por isso a variável foi desconsiderada no modelo final, pois não está alinhada com as expectativas de observações da realidade e com o consenso da literatura.

Novamente, com intuito de melhorar as possíveis inferências sobre a relação entre o preço do QAV e o ICMS, nova seleção de variáveis foi delimitada e optou-se pelo acréscimo da variável relacionada à uma aproximação para custos ‘avstl\_o’ (representa a média das etapas de voo do aeroporto de origem em km. Proxy para custos das cias aéreas, com variabilidade temporal, depende do número de cidades servidas).

Chega-se então ao modelo de preço do querosene de aviação, com todos os coeficientes estatisticamente relevantes, sem eliminação de variáveis e de sinal constante durante toda a análise de sensibilidade. Alguns pontos a serem considerados: (i) a relação positiva do ‘fuelpkm’ (representa média de litros consumidos de combustível por quilômetro voado nos voos da rota, conforme obtido do SINTAC) está alinhada com a hipótese de que quanto mais distante, mais caro o combustível devido aos custos de transporte até o local; (ii) a relação negativa entre ‘avstl\_o’ (média das etapas de voo da rota do aeroporto de origem em km) é outra validação já que ‘avstl\_o’ é uma aproximação para os custos da companhia aérea.

Percebe-se a grande relevância do módulo do coeficiente do preço do barril de petróleo (variável ‘wti\_o’) e também da alíquota do ICMS, representada pela variável ‘icms\_o’, ambos com sinal positivo indicando relação positiva, como já mencionado, ao aumentar essas variáveis é esperado aumento no preço do QAV (representado pelo regressando ‘fja1\_o’, em que o sufixo ‘o’ é indicativo de que o dado é referente à origem da rota).

Portanto, após seleção criteriosa das variáveis causais e investigativas, partindo de base ampla e com auxílio do LASSO e decisão manual foi possível construir um modelo em que se observa a influência do ICMS no preço do querosene de aviação.

Tabela 4: Regressão LASSO, nova seleção de variáveis e melhora nos resultados

	(1)	(2)	(3)	(4)	(5)	(6)
	lnfjal_o	lnfjal_o	lnfjal_o	lnfjal_o	lnfjal_o	lnfjal_o
lnwti_o	0.7284***	0.7284***	0.8679***	0.5737***	0.8682***	0.5736***
lnicms_o	0.1397***	0.1397***	0.1394***	0.1394***	0.1394***	0.1394***
lnavstl_o	-0.0828***	-0.0828***	-0.0098***	-0.0098***	-0.0094***	-0.0094***
lnfuelpkm	0.0401***	0.0401***	0.0069***	0.0069***	0.0060***	0.0060***
lnfuellt	0.0400***	0.0400***	0.0016***	0.0016***	0.0020***	0.0020***
lnschfltime					-0.0054***	-0.0054***
R2_Adj	0.7561	0.7561	0.9884	0.9884	0.9887	0.9887
RMSE	0.0899	0.0899	0.0087	0.0087	0.0087	0.0087
RMSE_CV	0.1030	0.1030	0.0225	0.0225	0.0223	0.0223
N_Obs	20,003	20,003	20,003	20,003	19,592	19,592
N_Vars	5	5	121	117	122	116
FE	yes	yes	yes	yes	yes	yes
Controls	no	no	yes	yes	yes	yes
Lasso	no	no	no	yes	no	yes

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

## 4.2 Modelo de Preço da Passagem Aérea

Para o modelo de Preço da Passagem Aérea o objetivo passa a ser a análise do preço da passagem aérea e sua relação com ICMS, para isso determinamos as variáveis causais, preço do QAV cheio e número total de voos (Regressores, variáveis explicativas) e vamos analisar as relações com o preço da passagem aérea (variável Regressando, Variável dependente, Variável de “desfecho”).

Verifica-se as relações entre as variáveis causais experimentadas durante a construção do modelo e as variáveis investigativas que estamos interessados, especialmente o ICMS. De posse dessas análises as variáveis são selecionadas novamente com inclusão e eliminação de variáveis ao longo do processo.

Para a nova seleção de variáveis os resultados que se mostram mais consistentes na análise de sensibilidade são mantidos. A relação entre QAV, sem adição do valor do ICMS, com o preço das passagens é esperada ser positiva e estatisticamente significativa ao longo de toda análise, assim como a média geométrica entre valores de origem e destino da alíquota do ICMS representada pela variável ‘gmicms’.

Nas análises para se chegar ao modelo final o LASSO indicou as melhores seleções de variáveis ao zerar coeficientes, além disso a autora optou por retirada manual devido a instabilidade do sinal das variáveis, como por exemplo ocorreu com o preço do combustível e

o número total de voos, além disso resultados que não estavam em consenso com a literatura também sofreram modificações.

Dessa forma, uma das variáveis retiradas manualmente foi a variável ‘fope’ (número total de voos operados, não cancelados), e também alterou-se a variável de QAV para ‘ja1’ que não inclui o valor do ICMS. Os resultados do modelo final após o processo de refinamento estão na Tabela 5.

Tabela 5: Regressão LASSO, ajuste de seleção de variáveis

	(1)	(2)	(3)	(4)	(5)	(6)
	lnp	lnp	Lnp	lnp	lnp	lnp
lnja1	0.2816***	0.2816***	0.9002***	0.9281***	0.9002***	0.9060***
lnf	-0.1510***	-0.1510***	-0.1206***	-0.1203***	-0.1206***	-0.1205***
lngmicms	0.0787***	0.0787***	0.0237**	0.0228**	0.0237**	0.0237**
lngmghi	0.5361***	0.5361***	0.3193***	0.3205***	0.3193***	0.3197***
fuelpkm			-0.0333***	-0.0330***	-0.0333***	-0.0333***
R2_Adj	0.1018	0.1018	0.1238	0.1235	0.1238	0.1239
RMSE	0.2284	0.2284	0.1970	0.1971	0.1970	0.1970
RMSE_CV	0.3834	0.3834	0.3950	0.3952	0.3952	0.3951
N_Obs	20,003	20,003	19,592	19,592	19,592	19,592
N_Vars	4	4	122	117	122	118
FE	yes	yes	yes	yes	yes	yes
Controls	no	no	yes	yes	yes	yes
Lasso	no	no	no	yes	no	yes

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Nesse modelo a inferência sobre a relação do preço das passagens e o ICMS é de uma relação positiva, ou seja, o aumento do ICMS provoca aumento no preço da passagem, assim como a diminuição do ICMS provoca queda no preço. A variável é estatisticamente significativa a 5% e tem magnitude relevante comparativamente às demais variáveis do modelo.

### 4.3 Modelo de Demanda

O modelo de demanda, como já ressaltamos na metodologia, possui características que o diferenciam dos outros dois modelos de preço, o objetivo é a análise da formação de demanda sem observação relacionada ao ICMS diretamente, porém é importante observar os fatores de formação da demanda separadamente, para poder dimensionar o impacto que outros acontecimentos, alheios às alterações nos impostos, podem gerar nos dois modelos de preço iniciais.

Para as análises iniciais das variáveis causais, vamos considerar a relação entre ‘rpax\_st’ (regressor, variável que está sendo construída por meio da regressão LASSO) e os principais

fatores que afetam a demanda: preço das passagens aéreas e disponibilidade de voos ou número total de voos no trecho observado (ambas serão as variáveis causais, regressandos).

Dos modelos elaborados inicialmente percebeu-se a instabilidade da variável preço, apesar da dimensão estar de acordo com o esperado o sinal da variável se altera na análise de sensibilidade, ou seja, isso demonstra que a confiabilidade dos resultados para o preço não é muito alta no modelo em questão. Portanto algumas alterações na seleção das variáveis foram propostas para eliminar essas variações de sinal e aperfeiçoar esses modelos iniciais de demanda.

A princípio, a análise utilizou duas variáveis relacionadas ao PIB, o que se verificou nos resultados ser um problema, por isso apenas uma delas, o PIB per capita ('gmgpc') foi mantida. A variável relacionada à população ('gmppop') foi retirada devido ao caráter genérico e não conclusivo que se observou na leitura dos resultados desses modelos. Foram inseridas na análise as variáveis 'maxconn' (representa o número máximo de passageiros em conexão), 'ncs' (representa o número de cidades servidas) e 'sseats\_chart' (representa o número de assentos disponíveis em voos charter).

Essas inserções tiveram como objetivos (i) observar o efeito dos passageiros em conexão, pois deveriam afetar o número total de passageiros em trânsito; (ii) observar o efeito do número de par de cidades servidas, pois isso demonstra a capilaridade do sistema e está em linha com a relação entre disponibilidade de voos e demanda e, por fim, (iii) observar o efeito da quantidade de assentos disponíveis em voos do tipo charter, pois são voos atrelados às atividades de turismo, uma variável que também nos remete à relação de disponibilidade e capilaridade da malha aérea, fatores essenciais na determinação da demanda.

Os resultados do modelo final, após o refinamento realizado pelo LASSO e por seleção manual por meio de análise cuidadosa e escolha criteriosa das variáveis estão na Tabela 6. Pode-se verificar que o coeficiente do preço das passagens, variável 'p', permanece negativo em toda a análise de sensibilidade do modelo de demanda, o que verifica a expectativa de relação negativa. Vale ressaltar a alta relevância do coeficiente, dado o módulo ser bastante alto comparado às outras variáveis do modelo.

Para o restante das análises percebe-se que o valor se torna positivo e com significância estatística a 1%, sendo então a leitura que será considerada neste trabalho.

Para as demais variáveis a expectativa se observa próxima do que é verificado nos resultados, de tal forma que 'f' (número total de voos), 'gmgpc' (PIB per capita), 'ncs' (número de cidades servidas) e 'sseats\_chart' (assentos disponíveis) são positivamente relacionadas a

demanda, ou seja o aumento dessas variáveis pode ocasionar aumento na demanda em magnitude semelhante ao módulo do coeficiente dessas variáveis.

Tabela 6: Regressão LASSO para Modelo de Demanda

	(1)	(2)	(3)	(4)	(5)	(6)
	Intickets	Intickets	Intickets	Intickets	Intickets	Intickets
lnp	-0.8240***	-0.8240***	-0.9211***	-0.9188***	-0.9211***	-0.9211***
lnf	0.2319***	0.2319***	0.1558***	0.1581***	0.1558***	0.1558***
lnfjal	0.3983***	0.3983***	0.1503**	0.1651**	0.1503**	0.1505**
lngmgpc	1.8217***	1.8217***	0.6542***	0.6776***	0.6542***	0.6540***
lnncs	0.2572***	0.2572***	0.0684***	0.0677***	0.0684***	0.0685***
lnsseats_chart	0.0100***	0.0100***	0.0173***	0.0173***	0.0173***	0.0173***
lnmaxconn	-0.1182***	-0.1182***	-0.0111	-0.0110	-0.0111	-0.0111
R2_Adj	0.6091	0.6091	0.5102	0.5143	0.5102	0.5102
RMSE	0.1942	0.1942	0.1402	0.1406	0.1402	0.1402
RMSE_CV	0.7242	0.7248	0.8570	0.8539	0.8566	0.8573
N_Obs	10,703	10,703	10,703	10,703	10,703	10,703
N_Vars	7	7	123	118	123	120
FE	yes	yes	yes	yes	yes	yes
Controls	no	no	yes	yes	yes	yes
Lasso	no	no	no	yes	no	yes

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Interessante observar que, como já mencionado, as variáveis ‘p’ (preço da passagem) são negativamente relacionadas com a demanda, e que ‘maxconn’ (número de passageiros em conexão) é estatisticamente não significativa, portanto não se pode fazer qualquer leitura do resultado e da relação dessa variável com a demanda, no caso o número total de passagens vendidas, diferente do que era esperado previamente neste estudo. Para as análises iniciais a relação foi negativa e estatisticamente significativa, o que é condizente com a realidade, dado que quanto mais conexões os passageiros tiverem em um mesmo voo menos tickets serão vendidos para a mesma quantidade de voos.

Outro ponto importante, é que a variável ‘maxconn’, que representa o número máximo de passageiros em conexão não tem significância estatística, porém não foi eliminada pelo LASSO. Vale ressaltar que o LASSO é um procedimento neutro, que analisa concomitantemente todos os valores, ou seja, a metodologia de refinamento pelo LASSO é uma abordagem multicritério qualitativa e argumentativa, precisa ser defendida por quem está modelando o problema, e nem sempre todos os critérios convergem.

De posse das análises individuais de cada modelo, vamos finalizar a discussão de resultados observando de forma conjunta o entendimento e sua relação com os objetivos deste trabalho.

#### 4.4 Discussão de Resultados

Dos resultados obtidos podemos validar o teste de hipótese, a relação entre o ICMS e o preço do querosene de aviação é verificada com um coeficiente na regressão LASSO positivo, estatisticamente significativa a 1%, com boa adequação dos dados demonstrado pelo  $R^2$ , que manteve valor maior que 0,5 em todos os casos de análise de sensibilidade.

Da mesma forma, a relação entre o preço do querosene de aviação e o preço da passagem nos oferecem dados que também comprovam a hipótese, pois o coeficiente de  $\beta_1$  na equação é positivo, estatisticamente significativa a 1% e, novamente, com boa adequação dos dados e dimensionalidade comparativamente às variáveis causais relevante.

A relevância do ICMS no setor de aviação civil foi testada no presente trabalho e mostrase evidências de que o aumento do ICMS gera aumento no preço do querosene de aviação, que por sua vez, hipótese também testada neste trabalho, vai ocasionar aumento no preço na passagem e diminuição na demanda. De posse do teste da hipótese, percebe-se que medidas que afetem a cobrança do ICMS vão gerar impactos significativos na indústria de aviação civil.

A recuperação do setor de transporte aéreo depende da colaboração do governo em manter esse serviço básico e uma das possíveis medidas colaborativas é o corte do ICMS, que como já vimos pode gerar aumento de demanda e, possivelmente, acelerar a recuperação do setor. Porém é válido mencionar que os níveis de emissão de gases diminuiu devido a crise causada pelo Covid-19 e que essa diminuição já poderia servir como referência para o setor e para o governo, que também é capaz de redirecionar o aumento do ICMS para a manutenção do nível de emissão de gases, que foi reduzido devido à crise.

Pelos coeficientes calculados, temos uma tabela com a dimensionalidade dos impactos de redução e aumento da alíquota do ICMS, já que todas as variáveis foram transformadas para logaritmo natural em todos os modelos: (i) preço do QAV; (ii) preço da passagem e (iii) demanda.

Tabela 7: Inferência sobre possíveis efeitos do ICMS no preço do QAV e no preço das passagens

Relação entre	Coeficiente da Relação	Aumento de 10% no ICMS	
		Efeito no Preço do QAV	Efeito no Preço da Passagem
Infja1 e icms	0,1394		
Inp e Infja1	0,906	1,4%	1,3%

Pela tabela acima temos que o aumento do ICMS em 10% pode causar aumento no preço da passagem de 1,3% aproximadamente, da mesma forma que a diminuição do imposto em 10% pode gerar diminuição do preço do QAV de 1,4%. Considerando o cenário nacional de aviação civil é passível de implementação medidas de redução do ICMS com intuito de colaborar na recuperação do setor. Novamente, esse aumento também poderia ser direcionado à taxaço de passageiro e de companhias aéreas com intuito de redução de demanda, e também de conscientizar a sociedade sobre seu papel na redução de emissão de gases e no desenvolvimento da indústria aeronáutica em prol dessa redução.

A evolução deste trabalho seria testar a hipótese e dimensionar as perdas de receita pelas companhias aéreas devido ao aumento do ICMS e contrapor esses resultados aos possíveis ganhos na área de preservação do meio ambiente. Vale também verificar se há aumento de pesquisas e investimento em desenvolvimento de aeronaves menos poluentes que vão, possivelmente, ajudar a diminuir a emissão.

## 5 Conclusão

O transporte aéreo é fundamental para um país de proporções continentais como Brasil, além de ser relevante para o cenário econômico do país, representando 1,4% do PIB, como mencionado no início deste trabalho. Assim como sua relevância econômica o transporte aéreo tem ganhado parcela significativa na emissão de gases poluentes, seu crescimento implica em aumento dessa parcela e, conseqüentemente, aumento de responsabilidade ambiental ao se utilizar esse tipo de transporte.

Construímos modelos que nos mostram evidências de que uma possível mitigação desses efeitos é uma decisão de posse do governo, que é capaz de aumentar ou reduzir a demanda por meio de tributação. Os modelos foram construídos por métodos de *machine learning*, com o auxílio da metodologia de regressão por encolhimento de variáveis e também por análises e observações retiradas desse método.

O fato da metodologia LASSO utilizada ainda estar em fase de desenvolvimento no Brasil, e no mundo, é uma dificuldade para o amplo entendimento da ferramenta devido à relativa escassez de trabalhos nacionais com essa aplicação, no entanto, como verificado nos resultados deste trabalho é uma ferramenta de extrema utilidade na inferência e entendimento de problemas reais como o abordado, sendo capaz de selecionar as melhores variáveis e, dessa forma, construir os melhores modelos. O LASSO possui implementação complexa, porém adaptável e de flexibilidade ímpar. Vale reforçar também a importância da base de dados NECTAR, muito bem construída e organizada, de forma a tornar o trabalho muito ágil, de alta confiabilidade e de fácil manipulação devido a sua documentação descritiva e explicativa de alta qualidade.

Com o auxílio da ferramenta LASSO e da base de dados o objetivo deste trabalho foi plenamente alcançado, já que podemos inferir com base nos resultados dos modelos de preço do QAV e preço das passagens a relação positiva entre o ICMS e ambos os regressandos. Além disso pela construção das variáveis em bases logaritmas podemos inferir e também modular essa relação e de posse desses valores formular interpretações desses resultados.

O coeficiente da relação entre o preço do querosene de aviação e a alíquota do ICMS foi 0,1394 para a regressão LASSO, indicando que cada aumento de 1% na alíquota do ICMS pode causar no preço do querosene de aviação um aumento de 0,1394%. Esse aumento resultará em aumento no preço da passagem, como pode ser inferido dos resultados do modelo de preço de passagens aéreas.

Nos modelos que construímos e analisamos neste trabalho o ICMS foi estatisticamente significativo na equação de preço nos dois modelos finais, pelos resultados deste trabalho há evidências na afirmação de que o ICMS afeta o preço das passagens aéreas. Apesar do grande impacto da pandemia no setor de aviação é importante retomar a discussão e direcionar esse aumento de impostos para redução de demanda e, conseqüentemente, redução na emissão de gases.

A redução do ICMS fomenta demanda, pois aumenta o número de voos e gera mais emissões de gases, bom pro transporte, mas não necessariamente bom pra questão do meio ambiente. É necessário observar os países que já estão discutindo a pauta há alguns anos, e decidiram mitigar as emissões por meio de tributação e, conseqüente, redução de demanda devido a esse aumento. A comparação deve servir apenas como referência, dado as idiosincrasias de cada região. A Europa possui outros meios de transporte que substituem o avião, como a malha ferroviária desenvolvida e de alta qualidade e velocidade, o que não se observa no Brasil, que possui matriz de transporte essencialmente rodoviária.

Os possíveis impactos ambientais da política tributária sobre o transporte aéreo gera um trade-off entre poluir e desenvolver de natureza complexa.

É notável que as considerações sobre o tema envolvem problemas relacionados às políticas nacionais e à economia do país, pois se trata de um setor fundamental para o desenvolvimento da sociedade e bem-estar da população, ainda mais diante de um cenário nacional de pandemia e de tentativa de recuperação econômica, em que o transporte aéreo ainda não apresenta sinais claros de recuperação. Deve-se buscar ao máximo otimizar as escolhas e planejar os possíveis desenlaces que as medidas de tributação podem gerar, sempre com o intuito de gerar benefícios de curto e longo prazo.

Possíveis desenvolvimentos deste trabalho seriam quantificar esse trade-off, de modo a oferecer ferramentas numéricas que auxiliem em decisões sobre esse trade-off e outro possível desenvolvimento é o questionamento sobre o desenvolvimento da malha aérea, essa seria a melhor opção para o Brasil ou há outros meios de transporte que podem contribuir mais e melhor para essa relação entre desenvolvimento e preservação do meio ambiente.

## 6 Bibliografia

1. Belloni, A., Chen, D., Chernozhukov, V., & Hansen, C. (2012). **Sparse models and methods for optimal instruments with an application to eminent domain.** *Econometrica*, 80(6), 2369-2429.
2. Belloni, A., Chernozhukov, V., & Hansen, C. (2014a). **Inference on treatment effects after selection among high-dimensional controls.** *The Review of Economic Studies*, 81(2), 608-650.
3. Belloni, A., Chernozhukov, V., & Hansen, C. (2014b). **High-dimensional methods and inference on structural and treatment effects.** *Journal of Economic Perspectives*, 28(2), 29-50.
4. Bendinelli, W. E. Bettini, H. F., & Oliveira, A. V. M. (2016). **Airline delays, congestion internalization and non-price spillover effects of low-cost carrier entry.** *Transportation Research Part A: Policy and Practice*, 85, 39-52.
5. Borsky, S., & Unterberger, C. (2019). **Bad weather and flight delays: The impact of sudden and slow onset weather events.** *Economics of Transportation*, 18, 10-26.
6. Chernozhukov, V., Hansen, C., & Spindler, M. (2015). **Post-selection and post-regularization inference in linear models with many controls and instruments.** *American Economic Review*, 105(5), 486-90.
7. Edwards, H. A., Dixon-Hardy, D., & Wadud, Z. (2016). **Aircraft cost index and the future of carbon emissions from air travel.** *Applied energy*, 164, 553-562.
8. Errico, A., & Di Vito, V (2017). **Performance-based navigation (PBN) with continuous descent operations (CDO) for efficient approach over highly protected zones,** *Transportation Research Part A: Policy and Practice*, 289-304.
9. Bilotkach, Volodymyr, & Lakew, Paulos Ashebir (2013). **On sources of market power in the airline industry: Panel data evidence from the US airports,** 24th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, 1-8.
10. Borenstein, Severin (1989). **Hubs and high fares: dominance and market power in the U.S. airline industry,** *RAND Journal of Economics*, Vol. 20, No. 3, 344-362.
11. Errico, A., & Di Vito, V (2017). **Performance-based navigation (PBN) with continuous descent operations (CDO) for efficient approach over highly protected zones,** 24th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, 1-8.