



UNIVERSIDADE FEDERAL DO RIO GRANDE
INSTITUTO DE CIÊNCIAS ECONÔMICAS, CONTÁBEIS E ADMINISTRATIVAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA APLICADA
MESTRADO EM ECONOMIA

THALIA IACKS CRIZEL

ALÉM DO CLICK: UMA EXPLORAÇÃO ECONÔMETRICA DA JORNADA DO
CONSUMIDOR VIRTUAL NO BRASIL

RIO GRANDE - RS

2024

Thalia lacks Crizel

Além do click: uma exploração econométrica da jornada do consumidor virtual no
Brasil

Dissertação apresentada ao Programa de Pós-Graduação em Economia Aplicada - PPGE da Universidade Federal do Rio Grande - FURG como requisito parcial à obtenção do grau de Mestre em Economia Aplicada.

Orientador: Prof. Dr. Cristiano Aguiar de Oliveira

Rio Grande - RS

2024

Ficha Catalográfica

C936a Crizel, Thalia lacks.

Além do *click*: uma exploração econométrica da jornada do consumidor virtual no Brasil / Thalia lacks Crizel. – 2024.

56 f.

Dissertação (mestrado) – Universidade Federal do Rio Grande – FURG, Programa de Pós-Graduação em Economia Aplicada, Rio Grande/RS, 2024.

Orientador: Dr. Cristiano Aguiar de Oliveira.

1. Comércio eletrônico 2. Perfil de Consumidores 3. Redes sociais
4. *Probit* 5. *Propensity score matching* 6. *Inverse Probability Weighted Regression Adjustment* I. Oliveira, Cristiano Aguiar de
II. Título.

CDU 339:004

Catálogo na Fonte: Bibliotecário José Paulo dos Santos CRB 10/2344



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO RIO GRANDE - FURG
COORDENAÇÃO DE CURSO DE MESTRADO EM
ECONOMIA APLICADA




ATA DE REUNIÃO, DE 27 DE JUNHO DE 2024

ATA DE DEFESA DE DISSERTAÇÃO


Aos 27 dias do mês de junho de dois mil e vinte e quatro, às 19h, realizou-se a defesa de dissertação do Programa de Pós-Graduação em Economia Aplicada, nível mestrado, da Universidade Federal do Rio Grande - FURG, no ambiente virtual da SEaD, cuja Banca foi constituída pelos professores: Cristiano Aguiar de Oliveira, Gibran da Silva Teixeira e Claudio Djissey Shikida, para arguir do mestranda **Thalia Iacks Crizel**. Após a apresentação da dissertação intitulada “**Além do click: uma exploração econométrica da jornada do consumidor virtual no Brasil**” e a arguição dos avaliadores seguida de defesa, a Banca reuniu-se e considerou o trabalho **APROVADO**, emitindo o parecer a seguir:

O trabalho atendeu a todos os requisitos para a sua aprovação, não sendo necessárias alterações relevantes.


Nada mais havendo a tratar, lavrou-se a presente Ata que após lida e aprovada será assinada pelos membros componentes da Banca.

Documento assinado digitalmente
 **CRISTIANO AGUIAR DE OLIVEIRA**
Data: 28/06/2024 10:18:13-0300
Verifique em <https://validar.iti.gov.br>

Prof. Cristiano Aguiar de Oliveira
Orientador

Documento assinado digitalmente
 **GIBRAN DA SILVA TEIXEIRA**
Data: 28/06/2024 14:28:18-0300
Verifique em <https://validar.iti.gov.br>

Prof. Gibran da Silva Teixeira
PPGE/FURG

Documento assinado digitalmente
 **CLAUDIO DJISSEY SHIKIDA**
Data: 28/06/2024 15:03:00-0300
Verifique em <https://validar.iti.gov.br>

Prof. Claudio Djissey Shikida
IBMEC/MG-PPGOM/UFPEL

RESUMO

O comércio eletrônico tem sido um agente transformador nas transações comerciais, especialmente nas últimas décadas, modificando significativamente a dinâmica do mercado. Este estudo objetiva examinar o impacto das características individuais e do uso das redes sociais na probabilidade de efetuar compras *on-line* no contexto brasileiro. Inicialmente, recorre-se a um modelo econométrico preditivo – *Probit* – para avaliar o impacto das características individuais na probabilidade de compra *on-line*. Com o propósito de investigar se existe uma relação causal entre o uso das redes sociais e o consumo *on-line*, utiliza-se o modelo econométrico *Propensity Score Matching* (PSM). A inclusão do modelo *Inverse Probability Weighted Regression Adjustment* (IPWRA) na metodologia aprimora o PSM, possibilitando um ajuste mais acurado das estimativas de efeito e tornando as inferências sobre o impacto do uso das redes sociais mais sólidas e confiáveis. A base de dados empregada provém do questionário da Pesquisa Tecnologia da Informação e Comunicação (TIC). As variáveis analisadas no modelo englobam sexo, idade, nível de escolaridade, renda familiar, etnia, *status* de atividade econômica, área de residência (rural ou urbana) e região geográfica. Os resultados indicam que várias características têm uma influência positiva e significativa na probabilidade de efetuar compras *on-line*, tais como viver em áreas urbanas, possuir um nível de escolaridade mais elevado, maior renda e ter o hábito de realizar transações financeiras pela internet. Ademais, infere-se que o uso das redes sociais aumenta a probabilidade de compra *on-line*.

Palavras-chave: Comércio eletrônico. Perfil de Consumidores. Redes sociais. *Probit*. *Propensity score matching*. *Inverse Probability Weighted Regression Adjustment*.

Classificação JEL: L81, M31.

ABSTRACT

E-commerce has been a transformative agent in commercial transactions, especially in recent decades, significantly changing the dynamics of the market. This study aims to examine the impact of individual characteristics and the use of social networks on the likelihood of making online purchases in the Brazilian context. Initially, a predictive econometric model - Probit - is used to assess the impact of individual characteristics on the likelihood of buying online. In order to investigate whether there is a causal relationship between the use of social networks and online consumption, the Propensity Score Matching (PSM) econometric model is used. The inclusion of the Inverse Probability Weighted Regression Adjustment (IPWRA) model in the methodology improves the PSM, enabling a more accurate adjustment of the effect estimates and making inferences about the impact of social media use more solid and reliable. The database used comes from the Pesquisa Tecnologia da Informação e Comunicação (TIC) questionnaire. The variables analyzed in the model include gender, age, level of education, family income, ethnicity, economic activity status, area of residence (rural or urban), and geographic region. The results indicate that several characteristics have a positive and significant influence on the likelihood of making online purchases, such as living in urban areas, having a higher level of education, having a higher income, and having the habit of making financial transactions online. Furthermore, it can be inferred that the use of social networks increases the likelihood of buying online.

Keywords: E-commerce. Consumer profile. Social networks. Probit. Propensity score matching. Inverse Probability Weighted Regression Adjustment.

JEL classification: L81, M31.

LISTA DE TABELAS

Tabela 1 - Estatísticas descritivas.....	21
Tabela 2 - Modelo <i>Probit</i>	30
Tabela 3 - PSM - Suporte comum.....	41
Tabela 4 - PSM - <i>Pstest</i>	42
Tabela 5 - Efeito do tratamento no grupo tratado sob diferentes métodos	43
Tabela 6 - <i>Probit</i> . Categoria de referência	54
Tabela 7 - Modelo <i>Propensity Score Matching</i> - <i>Logit</i>	55

LISTA DE GRÁFICOS

Gráfico 1 - Probabilidade de realizar compras <i>on-line</i> de acordo com o nível de escolaridade para cada gênero no Brasil	33
Gráfico 2 - Probabilidade de realizar compras <i>on-line</i> de acordo com o nível de renda familiar para cada gênero e idade	34
Gráfico 3 - Probabilidade de realizar compras <i>on-line</i> de acordo com o gênero, idade e <i>status</i> de atividade econômica - PEA.....	35
Gráfico 4 - Probabilidade de realizar compras <i>on-line</i> de acordo com o gênero, idade e etnia, no Brasil	36
Gráfico 5 - Probabilidade de realizar compras <i>on-line</i> de acordo com o gênero, idade e a área de residência.....	37
Gráfico 6 - Probabilidade de realizar compras <i>on-line</i> de acordo com o gênero, idade e região geográfica do Brasil.....	38
Gráfico 7 - Probabilidade de realizar compras <i>on-line</i> de acordo com o gênero, idade e utilização de redes sociais	39
Gráfico 8 - Probabilidade de realizar compras <i>on-line</i> de acordo com o gênero, idade e se realizou transações financeiras pela internet	40
Gráfico 9 - Suporte Comum	56
Gráfico 10 - PSM <i>Matching</i>	56

LISTA DE EQUAÇÕES

Equação 1	23
Equação 2	23
Equação 3	23
Equação 4	23
Equação 5	23
Equação 6	25
Equação 7	25
Equação 8	26
Equação 9	26
Equação 10	26
Equação 11	27

LISTA DE ABREVIATURAS E SIGLAS

ATET	<i>Average Treatment Effect on the Treated</i>
ATT	<i>Average Treatment Effect on the Treated</i>
CETIC.br	Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação da Unesco
CIA	<i>Conditional Independence Assumption</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
IPW	<i>Inverse Probability Weighted</i>
IPWRA	<i>Inverse Probability Weighted Regression Adjustment</i>
LR χ^2	Teste Qui-Quadrado de <i>Pearson</i>
LR	<i>Likelihood Ratio Test</i>
NIC.br	Núcleo de Informação e Coordenação do Ponto BR
PEA	População Economicamente Ativa
PNAD	Pesquisa Nacional de Amostras de Domicílios
PSM	<i>Propensity Score Matching</i>
TIC	Tecnologia da Informação e Comunicação

SUMÁRIO

1	INTRODUÇÃO.....	12
2	REFERENCIAL TEÓRICO	14
3	ESTRATÉGIA EMPÍRICA	20
3.1	Modelo <i>Probit</i>	22
3.2	Modelo <i>Propensity Score Matching</i>	24
4	RESULTADOS	29
4.1	<i>Probit</i>	29
4.2	<i>Propensity Score Matching (PSM)</i>	41
5	CONCLUSÕES.....	45
	REFERÊNCIAS	47
	APÊNDICE	54
	Apêndice A – <i>Probit</i>	54
	Apêndice B – <i>Propensity Score Matching (PSM)</i>	55

1 INTRODUÇÃO

O comércio eletrônico, também conhecido como *e-commerce*, tem experimentado um crescimento substancial nas últimas décadas, tornando-se parte fundamental da economia global contemporânea. É perceptível, em boa parte da população, a sobreposição deste veículo em detrimento dos meios tradicionais de comércio. O acesso generalizado às tecnologias digitais na vida cotidiana tem acelerado essa mudança, direcionando parcela significativa das transações comerciais para plataformas *on-line*.

Do ponto de vista dos negócios, a expansão do *e-commerce* tem suscitado interesse considerável para a exploração das potencialidades da internet em vias de otimizar as operações comerciais e financeiras. Nesse cenário, as estratégias de publicidade e *marketing* estão evoluindo consideravelmente, com o uso de algoritmos para direcionar produtos de forma mais precisa ao público-alvo, aquele com maior propensão à conversão (AGUIAR, 2023). Para que ocorra essa ligação entre o produto/serviço e o consumidor, é necessário conhecer o perfil deste.

Os economistas têm se concentrado no comportamento do consumidor *on-line*, na precificação dos bens e serviços digitais e nas características desse novo formato de comércio. Ainda, há interesses em comum entre economistas e estudiosos de *marketing*, no que tange a resposta do consumidor *on-line* a campanhas de *marketing*, publicidade e, principalmente, na capacidade das empresas em definir, segmentar mercados, públicos-alvo e posicionar produtos para maximizar retornos financeiros (LOPEZ, 2003).

Diversos fatores exercem influência na decisão de um indivíduo em realizar compras *on-line*, desde características pessoais até fatores socioeconômicos e culturais (CUNHA, 2021; NASCIMENTO, 2011). Ademais, o uso das redes sociais parece afetar diferentes aspectos sociais, psicológicos e comportamentais que, por sua vez, influenciam o comportamento de compra *on-line* dos usuários. Atualmente, as redes sociais desempenham um papel significativo como canal de comunicação entre as empresas e seu público-alvo, seja através de estratégias como tráfego pago, parcerias com influenciadores e campanhas publicitárias direcionadas a segmentos específicos de público (FERRARA, 2013; BARBOSA, 2010). Isto posto, a identificação dos perfis de consumidores e a respectiva ponderação desses fatores, bem como usar as redes sociais como ferramentas de *marketing* apresentam-se como importantes

aliados para a maximização do comércio eletrônico, setor de substancial importância para a economia nacional.

Para identificar os perfis dos consumidores com maior propensão ao consumo *on-line*, inicialmente, é feita uma contextualização do comércio eletrônico e sua interação com a literatura existente, serão explorados fatores que impulsionam e dificultam o desenvolvimento do *e-commerce*. Posteriormente, será feita uma análise da base de dados da Pesquisa Tecnologia da Informação e Comunicação (TIC) Domicílios, observando o comportamento dos consumidores e principais pontos de destaque dos questionários aplicados.

Após a revisão literária e análise dos dados, utiliza-se, inicialmente, como metodologia o modelo *Probit*, a fim de analisar como características individuais afetam a probabilidade de um indivíduo realizar compras *on-line*. Nesse aspecto, várias características indicam uma influência positiva e significativa na probabilidade de efetuar compras *on-line*, tais como viver em áreas urbanas, possuir um nível de escolaridade mais elevado, maior renda e ter o hábito de realizar transações financeiras pela internet.

Após essa análise, utiliza-se uma segunda metodologia, o modelo *Propensity Score Matching* (PSM), a fim de identificar se há relação causal entre o uso das redes sociais e o comércio eletrônico e, também, mitigar o viés de seleção para tornar as comparações entre os grupos mais plausíveis em termos de causalidade. Por fim, aplica-se o modelo *Inverse Probability Weighted Regression Adjustment* (IPWRA), para ajustar as estimativas de efeito ponderando às observações de acordo com as probabilidades inversas de seleção. A inclusão do IPWRA na metodologia contribui para melhorar a robustez das inferências causais, controlando ainda mais o viés de seleção e aprimorando a validade interna das estimativas do PSM. No âmbito do uso das redes sociais, infere-se que o uso das redes sociais aumenta a probabilidade de compra *on-line* de forma significativa.

O trabalho está organizado em mais quatro seções. A seção seguinte apresenta brevemente o comércio eletrônico no âmbito nacional, o papel das redes sociais no *e-commerce* e seu referencial teórico. A terceira seção apresenta os dados e a metodologia empregada. A quarta seção apresenta os resultados e suas respectivas discussões e, por fim, a quinta seção apresenta a conclusão da pesquisa.

2 REFERENCIAL TEÓRICO

O Brasil é um dos países com maior participação *on-line* no mundo. Uma pesquisa divulgada recentemente apontou que, em 2022, o Brasil contou com 149 milhões de usuários ativos na internet (CETIC.BR, 2023). Alia-se a essa predisposição às redes a ampliação do acesso, por parte da população, aos aparelhos com acesso à internet, como aponta a Pesquisa Nacional por Amostras de Domicílios (PNAD), quando o acesso à internet ultrapassou a marca de 90% dos lares brasileiros (IBGE, 2021).

O varejo brasileiro pode ser dividido em dois grupos, um composto por inúmeras empresas pequenas com atuação local e regional, e outro com poucas grandes empresas com atuação nacional como Americanas, Submarino, Magazine Luiza¹, entre outras. Nesse contexto, a literatura demonstra que ambas categorias podem se beneficiar pelo processo de digitalização (IPEA, 2011).

Em conjunto com o crescimento do acesso à internet, mudam-se os hábitos cotidianos e cresce a busca da população por bens e serviços nos meios digitais, sobretudo pela comodidade e economia proporcionadas pela tecnologia, permitindo um ambiente em que é possível comparar preços e produtos de forma muito mais rápida e dinâmica do que no mercado tradicional (IPEA, 2011). O novo panorama aflorou o senso de urgência das empresas para se manterem competitivas adequando suas transações comerciais às formas de comércio eletrônico (FARIA; CARVALHO; CARVALHO, 2012).

No entanto, a despeito do crescimento constante do acesso e tempo de permanência dos usuários na internet e, por conseguinte, do aumento da exposição às publicidades em *sites*, redes sociais e outros aplicativos, historicamente a taxa de conversão de usuários visitantes em compradores é considerada baixa (BAYMARD INSTITUTE, 2023; MORGADO, 2003). Em uma pesquisa recente, Zhang et al. (2017) investigaram a influência das redes sociais no comportamento de consumo dos usuários, e descobriram que a utilização cumulativa de redes sociais está positivamente associada à atividade de compras. No entanto, também se observa uma relação negativa de curto prazo: logo após um período de maior utilização das redes sociais, a atividade de compras *on-line* tende a ser menor. Este duplo efeito

¹ Para maiores *insights* sobre a atuação do *e-commerce* da Magazine Luiza ver: Balestrin e Darigo (2014).

indica que, embora as redes sociais possam levar ao aumento da descoberta e do interesse nos produtos ao longo do tempo, o uso intenso e imediato pode desviar a atenção das atividades reais de compras *on-line*. Diante disso, as empresas e estudiosos de *marketing* buscam estratégias para reverter a visita virtual em vendas efetivas. E neste ponto, temos dois fatores que contribuem de forma significativa para essa conversão: (1) a identificação do perfil do consumidor *on-line* e (2) o uso das redes sociais.

Nesse cenário, a identificação do perfil dos consumidores se faz necessária, permitindo que as empresas compreendam o tamanho do mercado e sua setorização, com intuito de gerar tráfego *on-line* qualificado, aproximando os potenciais compradores – até então meramente usuários da internet – dos produtos e serviços compatíveis com suas necessidades e perfil, potencializando a probabilidade de compra.

Podemos destacar algumas pesquisas que realizaram essa identificação. O estudo do IPEA (2011) descreve o perfil do consumidor *on-line* brasileiro no ano de 2011. Contudo, a pesquisa não evidencia significância estatística entre o uso de redes sociais e o comércio eletrônico. A tese de Morgado (2003) traça o perfil do consumidor *on-line* brasileiro, identificando comportamentos e preferências desses consumidores na internet. Outras pesquisas destacam que características como renda e escolaridade superior estão positivamente relacionadas com a atividade de compra *on-line*, tais como a já referida pesquisa de Morgado (2003) e outras, como Song e Sun (2020), Lima Filho et al. (2012) e Li, Kuo e Russell (1999). Na pesquisa de Song e Sun (2020), indivíduos com rendimentos mais elevados, maior nível de escolaridade, jovens, mulheres, casados, residentes urbanos e que utilizam smartphones são mais propensos a realizarem compras *on-line* na China.

O impacto do uso das redes sociais no consumo *on-line* é uma questão multifacetada, abrangendo vários aspectos do comportamento do usuário, desde efeitos psicológicos a mudanças nos padrões de consumo. Uma pesquisa recente buscou analisar o impacto das redes sociais no desempenho acadêmico dos alunos, e verificou-se uma relação negativa e significativa, quanto mais tempo nas redes sociais, menor tende a ser o desempenho acadêmico dos alunos (PAUL; BAKER; COCHRAN, 2012).

Há também efeitos sociais e psicológicos envolvidos no uso das redes sociais, como demonstrado por Wilcox e Stephen (2012), os usuários focados em amigos

próximos, enquanto navegam em suas redes sociais, experimentam um aumento momentâneo na autoestima, o que reduz o autocontrole após navegar em uma rede social. Além disso, o consumo frequente das redes sociais está ligado ao aumento da procrastinação e à menor satisfação com a vida (HINSCH; SHELDON, 2013). Portanto o uso das redes sociais gera uma preocupação com a dependência dessas ferramentas, que pode levar a distúrbios psicológicos em algumas pessoas. A situação foi agravada pela pandemia de COVID-19, que gerou um elevado consumo de redes sociais e a incidências significativas de dependência (GÓMEZ-GALÁN et al., 2020).

A influência do uso das redes sociais no comportamento de consumo é uma área de grande interesse para economistas, profissionais de *marketing* e consumidores. Um estudo recente, realizado por Shi e Wang (2020), examinou como as redes sociais afetam o comportamento de consumo de famílias chinesas, focando especificamente na restrição de crédito. Ao analisar os microdados de 27.632 famílias, os pesquisadores descobriram que a "cultura de grupo" *on-line*, uma forma de capital social emergente na China, influencia significativamente o crédito e os gastos dos consumidores. Isso ressalta como as redes sociais, ao promoverem esse novo tipo de capital social, podem moldar os padrões de consumo das famílias, especialmente no que diz respeito ao uso do crédito, e têm implicações importantes para entender o impacto econômico mais amplo das redes sociais no consumo.

Um estudo realizado no Paquistão, em 2015, consistiu em um questionário aplicado a uma amostra de 1.000 indivíduos, com faixa etária de 18 a 50 anos, e concluiu que as redes sociais apresentam um impacto significativo no comportamento de consumo *on-line* no país (MADNI, 2014). Ainda, outra pesquisa realizada especificamente para o setor de vestuário, explora a relação entre a comunicação nas redes sociais, a experiência de compra *on-line* e a pesquisa pré-compra, através de questionários, dos quais obtiveram 458 respostas, e concluiu-se que as redes sociais são a plataforma-chave para alcançar os clientes (SAHA; SAHNEY, 2021). Esses resultados estão de acordo com o ambiente interativo gerado pelas redes sociais, no qual as marcas podem se envolver diretamente com os consumidores. Isso cria um senso de comunidade em torno da marca e permite que as empresas construam relacionamentos mais próximos com seu público-alvo. O engajamento por meio de curtidas, comentários, compartilhamentos e mensagens diretas permite uma

comunicação bidirecional, na qual as empresas podem obter *feedback* instantâneo dos clientes e responder às suas perguntas e preocupações em tempo real.

Ainda no contexto das redes sociais, surgem os *digitais influencers*² e os *streamers*³, que atuam como uma ponte de aproximação entre as empresas e os clientes, auxiliando no processo de conversão das vendas *on-line*, suprimindo algumas falhas do comércio eletrônico, a exemplo da experimentação dos produtos. Em geral, indivíduos que compram *on-line* têm orientação experiencial menor (MORGADO, 2003), entretanto há uma parcela expressiva de usuários que deixam de realizar compras *on-line* devido ao apeço a experiência, nesse aspecto, quando comparado ao comércio tradicional, o comércio eletrônico está em desvantagem (ANDRADE; SILVA, 2017).

Os influenciadores digitais através das “*lives*” e conteúdos publicitários, promovem uma experiência “quase real” do produto, e ainda dispõem de um fator adicional: a confiança do usuário. A literatura demonstra que os influenciadores possuem uma relação de proximidade com o seu público, que por sua vez, confia no conteúdo desses profissionais, assim a confiança é um fator importante e está associada de forma positiva com a decisão de realizar compras *on-line* (HUANG, 2021; WU; ZHANG et al, 2023; HAJLI, 2014; UZUNOĞLU; KIP, 2014; ZHANG et al, 2022). Ainda, um estudo realizado na China, examinou os fatores que influenciam as intenções de compras contínuas dos consumidores no TikTok®, plataforma responsável por mais da metade das compras por transmissão ao vivo na China, os resultados destacam a importância da experiência do usuário e da satisfação em relação ao conteúdo na construção de intenções de compras duradouras (WANG; OH, 2023).

Concluindo, o impacto do uso das redes sociais no consumo e no comportamento *on-line* é complexo, com efeitos positivos e negativos. Embora as redes sociais possam melhorar a atividade de compra e aumentar as taxas de conversão de vendas *on-line*, também apresentam riscos de dependência, diminuem a satisfação com a vida e têm impactos negativos no desempenho acadêmico e no

² De acordo com Salgado (2022), do blog Opinion Box, *influencers* digitais são profissionais capazes de influenciar e formar a opinião de outras pessoas, por meio de conteúdos que elas produzem nas redes sociais.

³ No *blog* Netshow.me, Arcoverde (2023), descreve o *streamer* como um profissional que trabalha com a criação de vídeos *on-line*, utilizando diversas plataformas para se conectar com seu público, ampliar sua rede, fechar parcerias com marcas e ganhar porcentagens de anúncios.

autocontrole. Compreender estes efeitos pode ajudar os usuários a navegar nas redes sociais de forma mais eficaz e a mitigar potenciais consequências negativas.

A estruturação da internet permite aos administradores de *sites*, aplicativos e provedores gerarem uma grande quantidade de dados acerca do comportamento dos usuários, desde o tipo de conteúdo que consomem por mais ou menos tempo, até o tipo de publicidade que possuem maior predisposição de consumo. Para Shapiro e Varian (1999), as empresas que forem as primeiras e melhores em utilizar as informações para conhecer seus clientes potenciais terão vantagens competitivas e maior probabilidade de maximização de lucros em comparação com as demais.

Nesta pesquisa, conforme veremos detalhadamente na seção de Estratégia Empírica, utiliza-se a Pesquisa TIC Domicílios como base de dados, de acordo com ela o Brasil contava com 149 milhões de usuários de internet, equivalente a 81% da população. Destes, aproximadamente 45% fizeram uso do comércio eletrônico no período de 12 meses, alcançando o patamar de 67 milhões de consumidores no ano de 2022 (NIC.br, 2023).

A literatura trata de alguns pontos que influenciam o consumo *on-line*, como a acessibilidade por parte dos indivíduos a aparelhos e plataformas de acesso à rede, a segurança das plataformas contra possíveis fraudes, infraestrutura e logística de entrega, bem como a pandemia de COVID-19 e o consequente distanciamento social.

Acerca do crescimento da obtenção, por parte dos indivíduos, de aparelhos com acesso à internet, as pesquisas apontam para uma relação positiva (MORGADO, 2003), tendo em vista que nos últimos anos, conforme mencionado anteriormente, o Brasil alcançou patamares históricos no que diz respeito ao percentual de usuários ativos, em conjunto com o crescimento, também, do consumo *on-line* (IPEA, 2011).

A segurança das plataformas digitais é outro fator que influencia substancialmente a decisão de comprar *on-line* em detrimento de meios tradicionais de consumo (MORGADO, 2003). O critério “segurança” diz respeito desde à segurança dos dados dos usuários, pelo vazamento de informações pessoais (inclusive dados bancários) até dúvidas acerca do recebimento do produto em si e a certificação do vendedor. Nesse aspecto o comércio tradicional se sobrepõe ao *e-commerce* (ANDRADE; SILVA, 2017).

O desenvolvimento da infraestrutura nacional para escoamento de mercadorias, bem como o aprimoramento das políticas de logística empregadas pelas empresas varejistas e locais, abreviaram os prazos de entrega e aproximaram as

plataformas digitais dos consumidores mais imediatistas. Esse processo modificou a relação entre consumidor e vendedor/prestador de serviço, permitindo o surgimento de novos modelos de negócio, a exemplo das *Dark Kitchens*⁴, Uber®, iFood®, entre outros.

Por fim, este cenário sofreu substancial impacto durante a pandemia de COVID-19, onde a digitalização de serviços e a oferta de produtos em *sites* e aplicativos apresentou grande crescimento, em razão do necessário distanciamento social, conforme observado por Cardoso, Kawamoto e Massuda (2019). Essa aceleração do processo de digitalização de vários setores econômicos trouxe melhorias para o ambiente virtual e popularizou o *e-commerce* para algumas camadas de consumidores que ainda não haviam experimentado este comércio (FAVORETE; PEREIRA, 2021; PREMEBIDA, 2021; SILVA et al., 2021).

Na seção seguinte, será abordada a estratégia empírica utilizada. Inicialmente, utiliza-se o *Probit* a fim de prever e compreender como as características individuais afetam a probabilidade de consumo *on-line* no Brasil. Após, utiliza-se o PSM para uma análise de causalidade entre o uso das redes sociais e o consumo *on-line*. Por fim, como complemento, verifica-se a robustez do modelo através do método IPWRA.

⁴ O termo *Dark Kitchens* refere-se a um modelo de negócios que advém da evolução do mercado de alimentos que combina duas transformações tecnológicas disruptivas: o uso de plataformas digitais e de estruturas de produção e logística que permitem reduzir custos para vendedores e aumentar a satisfação dos consumidores. Para maiores informações ver Oliveira (2023).

3 ESTRATÉGIA EMPÍRICA

A metodologia empregada na pesquisa utiliza o modelo *Probit*, seguido do PSM e IPWRA. Antes de detalhar cada modelo, faz-se necessário analisar a base de dados e as variáveis construídas a partir da revisão de literatura e disponibilidade de informações, bem como definir algumas limitações empregadas para melhor adaptação da pesquisa.

A base de dados utilizada advém da Pesquisa TIC Domicílios, com dados agrupados para todos os anos disponíveis, sendo 2015 o primeiro ano em que a pesquisa foi realizada, seguindo até o ano de 2022. Cabe destacar que no ano de 2020 não foram realizadas perguntas do módulo de comércio eletrônico (módulo H), portanto, esse ano não foi incluído na pesquisa⁵. Finalmente, os dados agrupados são de 2015, 2016, 2017, 2018, 2019, 2021 e 2022⁶. A pesquisa conta com vários indicadores para mapear o cenário do acesso e do uso das tecnologias de informação e comunicação no Brasil, sendo um dos temas abordados o comércio eletrônico.

A variável dependente é construída a partir da pergunta H2, que questiona: “O (a) Sr. (a) comprou ou encomendou produtos ou serviços pela Internet nos últimos 12 meses, mesmo que o pagamento não tenha sido feito pela Internet?”. Ou seja, a pergunta identifica se o indivíduo participou do comércio eletrônico ou não. Seguindo alguns estudos empíricos que avaliam o perfil do consumidor *on-line* (IPEA, 2011; MORGADO, 2003; LOHSE; BELLMAN; JOHNSON, 2000), as variáveis independentes são: (i) características físicas, geográficas e socioeconômicas: sexo, idade, escolaridade, renda familiar, etnia, região de residência, área urbana ou rural, *status* de atividade econômica (PEA⁷); (ii) comportamentais: se o indivíduo utiliza redes sociais e se realiza transações financeiras pela internet. A Tabela 1 fornece estatísticas descritivas das variáveis utilizadas.

⁵ Para maiores informações acesse CETIC.br (2021).

⁶ Para obter os microdados utilizados na pesquisa, acesse: NIC.br, 2015; NIC.br, 2016; NIC.br, 2017; NIC.br, 2018; NIC.br, 2019; NIC.br, 2021; NIC.br, 2022.

⁷ População Economicamente Ativa.

Tabela 1 - Estatísticas descritivas

	Mulher	Homem	Total
Comprou <i>on-line</i>	33,43	39,67	36,27
Gênero	54,55	45,45	100
Idade	32,97	31,95	32,51
Trabalha (PEA)	70,73	83,84	76,69
Área urbana	90,14	90,57	90,33
Redes sociais	81,28	76,68	79,19
Trans. financeira	30,03	37,03	33,21
Escolaridade			
Analfabeto/Educação Infantil	2,65	3,01	2,85
Fundamental	29,64	32,21	30,81
Médio	48,93	47,91	48,46
Superior	18,78	16,87	17,91
Renda familiar			
Até 1 SM	40,78	30,00	35,88
Mais de 1 SM até 2 SM	27,06	26,98	27,02
Mais de 2 SM até 3 SM	15,24	18,49	16,72
Mais de 3 SM até 5 SM	10,73	14,03	12,23
Mais de 5 SM	6,19	10,50	8,15
Etnia			
Branca	32,06	32,63	32,32
Preta	14,23	15,85	14,96
Parda	48,94	47,06	48,09
Amarela	3,30	2,27	2,83
Indígena	1,48	2,19	1,80
Região			
Sudeste	28,31	31,24	29,64
Nordeste	29,40	27,70	28,63
Sul	14,10	13,97	14,04
Norte	17,56	16,69	17,16
Centro-Oeste	10,63	10,39	10,52

Fonte: elaboração própria.

Nota: exceto a idade, todas as outras variáveis estão em percentagens da amostra.

As estatísticas descritivas mostram que mais de 39% dos homens relatam ter realizado compras *on-line* em comparação com 33% das mulheres, essa distinção – consumo do sexo masculino superior ao feminino – também é observada em outras pesquisas (IPEA, 2011; MORGADO, 2003). Além disso, a partir das estatísticas descritivas, é possível inferir que o entrevistado médio vive em áreas urbanas, tem em torno de 32 anos, está trabalhando, utiliza redes sociais, concluiu o ensino médio, tem renda de até um salário mínimo e se declara pardo.

O acesso às tecnologias e ao comércio eletrônico é afetado pela idade dos indivíduos (SOUBATI, 2009). A questão aqui é o ponto de ruptura entre indivíduos que estão ativos na internet e os que não estão. Neste estudo optou-se por limitar a

análise a indivíduos com idade igual ou inferior a 60 anos, esse ponto de ruptura é estabelecido analisando a base de dados em questão. Indivíduos com mais de 60 anos apresentam pouca ou nenhuma interação com as tecnologias de informação e comunicação e temos uma menor prevalência de compras *on-line* depois dos 60 anos, possivelmente por questões de acessibilidade e familiaridade com a internet.

Entretanto, esse ponto de ruptura é dinâmico. Estudos recentes indicam que, com o avanço das tecnologias, as pessoas com mais de 75 anos (ponto de ruptura) hoje em dia são mais ativas do que as que tinham a mesma idade há dez anos. Esse processo ocorre devido à crescente exposição e incorporação da tecnologia na sociedade e na economia⁸.

Por outro lado, o limite inferior da amostra, em termos de idade, é de 10 anos⁹. Considerando que o acesso às ferramentas tecnológicas ocorre de forma precoce, torna-se viável incluir indivíduos de menoridade. Cabe destacar que, por esse motivo, utilizamos no modelo a renda familiar, a fim de captar a influência dos menores na decisão de compra das famílias. Com essas restrições, a amostra contém 82.339 indivíduos, sendo 44.920 mulheres (54,55% do total) e 37.419 homens (45,45% do total).

3.1 Modelo *Probit*

Determinar características que afetem a probabilidade de realizar compras *on-line* não é trivial e, como já explorado, esse não é um objetivo de estudo exclusivo da economia, outras áreas de pesquisa, a exemplo do *marketing* digital, têm interesse em determinar o perfil do consumidor *on-line*. Seguindo a literatura econômica, é possível analisar a decisão de comprar *on-line* através de uma análise de custo-benefício, modelada como uma função de utilidade aleatória em que os indivíduos realizam compras *on-line* se a utilidade de comprar (c) for maior do que a utilidade de não comprar (nc), dadas respectivamente por:

⁸ Para maiores *insights* acerca da relação entre a idade e o comércio eletrônico, ver Rybaczewska e Sparks (2022).

⁹ No modelo *Probit* utiliza-se uma amostra com indivíduos com idade entre 10 e 60 anos. No PSM utiliza-se uma amostra de indivíduos com idade entre 18 e 60 anos devido ao suporte comum (para melhor pareamento dos grupos).

Equação 1

$$U_c = x'_c \beta_c + \varepsilon_c$$

Equação 2

$$U_{nc} = x'_{nc} \beta_{nc} + \varepsilon_{nc}$$

Em que x_c e x_{nc} representam um conjunto de características que afetam a utilidade de realizar compras *on-line*, e ε_c e ε_{nc} representam os componentes aleatórios (não observáveis) da função de utilidade. Portanto, se $x'_c \beta_c + \varepsilon_c \geq x'_{nc} \beta_{nc} + \varepsilon_{nc}$, é racional realizar compras *on-line* e, nesse sentido, a probabilidade de comprar é igual a um. Suponha que $\varepsilon_c - \varepsilon_{nc} = \varepsilon \sim F_\varepsilon$, onde F é simétrico. Denote $x' \beta = x'_c \beta_c - x'_{nc} \beta_{nc}$, de modo que:

Equação 3

$$\Pr\left(\text{comprou} = \frac{1}{x}\right) = \Pr\left(x' \beta + \varepsilon \geq \frac{0}{x}\right) = \Pr\left(\varepsilon \geq -\frac{x' \beta}{x}\right)$$

Portanto, por simetria:

Equação 4

$$\Pr\left(\text{comprou} = \frac{1}{x}\right) = \Pr\left(\varepsilon \leq \frac{x' \beta}{x}\right) = F_\varepsilon(x' \beta)$$

Essa probabilidade condicional pode ser obtida, por exemplo, por meio de um modelo *Probit*, ou seja, se $\varepsilon_c, \varepsilon_{nc} \sim N(0, \Sigma)$ e $\varepsilon \sim N(0, 1)$. Assim, a probabilidade de um indivíduo comprar *on-line* condicionadas às características definidas no modelo teórico é dada por:

Equação 5

$$\Pr\left(\text{comprou} = \frac{1}{x}\right) = \Phi(x' \beta) = \int_{-\infty}^{x' \beta} \frac{1}{\sqrt{2\pi}} e^{-1/2 t^2} dt$$

Essa metodologia permite identificar as características que impactam, seja de forma positiva ou negativa, a probabilidade dos indivíduos de utilizarem o comércio

eletrônico. Entretanto, nada se pode inferir sobre causalidade até então. Portanto, na subseção seguinte, utiliza-se uma outra estratégia empírica, o modelo PSM, que permite tratar o possível problema de viés de seleção e, também, gerar grupos factíveis e comparáveis.

3.2 Modelo *Propensity Score Matching*

As redes sociais desempenham um papel fundamental no comércio eletrônico, conforme já abordado, oferecendo uma plataforma poderosa para as marcas se conectarem com os consumidores, promoverem seus produtos e serviços e impulsionarem vendas. Portanto, a pesquisa se propõe a avaliar o impacto da utilização de redes sociais sobre a probabilidade de indivíduos realizarem compras *on-line*. Neste cenário, tendo em vista que a natureza da amostra impede a análise simultânea de um mesmo indivíduo enquadrado como tratamento (utiliza rede social) e controle (não utiliza rede social), torna-se necessária a aplicação de ferramentas econométricas para realizar a análise.

A ferramenta inicialmente eleita para a realização desta análise é o PSM (ROSENBAUN; RUBIN, 1983), que visa a verificação de relação causal entre objetos. Para isso a ferramenta é utilizada com o intuito de reduzir ou eliminar o viés de seleção das variáveis observáveis e tornar os grupos da amostra estatisticamente comparáveis, um exemplo dessa aplicação é a pesquisa de Markus, Oliveira e Nishimura (2020).

Este modelo pressupõe a comparação entre grupos de indivíduos tratados (que utilizam redes sociais) e não tratados (que não utilizam), para tanto, é fundamental que os grupos sejam passíveis de comparação, isto é, que as características observáveis sejam similares, reduzindo, por este fim, o viés do modelo. O modelo *logit* é empregado para criar esse grupo de comparação. Ainda, as funções do *logit* ou *probit*, no modelo PSM, servem para calcular o escore de propensão, ou seja, a probabilidade de um indivíduo receber o tratamento, considerando suas características observáveis. Nesta pesquisa utiliza-se o modelo *logit*, e o tratamento, como já mencionado, é ter utilizado redes sociais.

Assim, formalmente temos: a) seja “T” a variável que representa o tratamento, assumindo valor 1 caso o indivíduo tenha utilizado rede social e 0 caso contrário; b) seja $Y_i(1)$ os resultados potenciais do indivíduo i , caso tenha utilizado rede social e feito compras *on-line* e $Y_i(0)$ o resultado potencial para quem não recebe o tratamento,

mas comprou *on-line*. E, ainda, c) seja X_i o vetor das variáveis explicativas, composto pelas características individuais observáveis já detalhadas no início desta seção.

Para se obter a estimação do efeito causal de utilizar redes sociais sobre a probabilidade de realizar compras *on-line*, calcula-se o efeito médio do tratamento sobre os tratados – ATT (*average treatment effect on the treated*), por meio da diferença da média entre tratados e controles, a saber:

Equação 6

$$ATT(x) = E[Y_i(1)|T_i = 1, X_i = x] - E[Y_i(0)|T_i = 1, X_i = x]$$

A expressão $E[Y_i(0)|T_i = 1, X_i = x]$, na Equação 6, indica a média que os indivíduos tratados teriam se não tivessem recebido o tratamento (ou seja, a média dos tratados caso fossem controles), com base em suas características observadas. No entanto, esse valor não pode ser diretamente observado. Portanto, é necessário encontrar um substituto apropriado para permitir a estimativa do parâmetro ATT. Para isso, o modelo adota uma estratégia de identificação na qual, dadas as características observáveis X_i , que incluem aspectos pessoais, socioeconômicos e culturais dos indivíduos, os resultados potenciais seriam independentes do tratamento.

Essa é a premissa da independência condicional (CIA – *Conditional Independence Assumption*), proposta por Rosenbaum e Rubin (1983). Segundo ela, o conjunto de variáveis observadas X_i deve conter todas as informações sobre o resultado potencial na ausência do tratamento – $Y_i(0)$ – que o indivíduo possui ao decidir participar ou não do tratamento. Quando controlamos pelo vetor X_i (que é composto por variáveis observadas), as variáveis $Y_i(0)$ e $Y_i(1)$ tornam-se independentes da variável binária de tratamento T_i . Temos assim:

Equação 7

$$Y_i(1), Y_i(0) \perp T_i | X_i$$

Isso significa que, quando consideramos as características individuais, X_i , os resultados potenciais não estão relacionados ao uso de redes sociais. Isso implica que o resultado de um indivíduo no grupo de controle pode ser um bom indicador do resultado potencial de um indivíduo no grupo de tratamento, na ausência de tratamento, desde que possuam o mesmo conjunto de variáveis observadas. Portanto,

de acordo com a premissa da independência condicional, todas as variáveis que influenciam o tratamento (uso de redes sociais) e/ou os resultados devem ser controladas pelo modelo.

Com base nessa suposição, podemos reformular o segundo termo da Equação 6, de modo que o ATT dependa apenas de variáveis observáveis, possibilitando assim seu cálculo. Portanto, temos:

Equação 8

$$ATT(x) = E[Y_i(1)|T_i = 1, X_i = x] - E[Y_i(0)|T_i = 0, X = x]$$

Outra suposição essencial para o modelo é a hipótese de sobreposição, também conhecida como suporte comum. Essa hipótese assegura que cada indivíduo no grupo de tratamento possua um correspondente próximo no grupo de controle, no qual a variável resultado refletiria a situação desse indivíduo na ausência do tratamento, e vice-versa. Essa hipótese pode ser formulada como:

Equação 9

$$0 < \Pr[T_i = 1|X_i] < 1$$

Antes de implementar o PSM, é necessário calcular o escore de propensão, já que a probabilidade de um indivíduo receber o tratamento – considerando suas características observáveis X_i – não é conhecida. Para essa estimativa, recorre-se ao modelo paramétrico *logit*¹⁰. De acordo com Rosenbaum e Rubin (1983), o escore de propensão é definido como:

Equação 10

$$P(X) = \text{Probabilidade}[T = 1|X]$$

Por último, levando em consideração as suposições de suporte comum e independência condicional, juntamente com os escores de propensão, o efeito médio do tratamento sobre os tratados – ATT – pode ser calculado como:

¹⁰ Ainda, temos que a probabilidade de participar ou não segue o seguinte modelo:

$$\Pr[T_i = 1|X_i = x] = \frac{\exp(x\beta)}{1 + \exp(x\beta)}$$

Equação 11

$$ATT(x) = E[Y_i(1)|T_i = 1, P(X)] - E[Y_i(0)|T_i = 0, P(X)]$$

Adaptando as equações ao modelo, temos que as características observáveis utilizadas seguem as mesmas do modelo *Probit*, portanto a Equação 12 explica a probabilidade de compra *on-line*, por meio da utilização de redes sociais, dadas as características observadas.

Equação 12

$$\begin{aligned} Comprou = & \beta_0 + \beta_1 ano + \beta_2 sexo + \beta_3 pea + \beta_4 renda familiar + \beta_5 escolaridade \\ & + \beta_6 idade + \beta_7 idade^2 + \beta_8 etnia + \beta_9 área + \beta_{10} região + \varepsilon_i \end{aligned}$$

Ainda, aplica-se alguns filtros para a equação estimada, o primeiro limita a idade dos indivíduos analisados, entre 18 e 60 anos, e também filtramos por indivíduos que responderam “Sim” para a pergunta c5_d: “Nos últimos 3 meses, o respondente utilizou a Internet no telefone celular?”. Esse último filtro serve para separar os indivíduos que realmente utilizam internet de forma regular dos que utilizam de forma esporádica.

É importante destacar que o modelo adotado apresenta limitações, pois, de acordo com a hipótese de independência condicional, todas as variáveis que impactam o tratamento e/ou os resultados devem ser controladas no modelo (ROSENBAUM; RUBIN, 1983), o que é bastante discutível. Além disso, características individuais não observáveis (*confunders*) podem influenciar o tipo de tratamento que um indivíduo recebe, possivelmente afetando o efeito do uso de redes sociais. Em face disso, utiliza-se métodos para análise de robustez do modelo.

Após obter os escores de propensão, faz-se necessário determinar as técnicas de pareamento que serão utilizadas. De acordo com Rubin (2001), as técnicas de pareamento podem ser classificadas em três tipos: correspondência, subclassificação e ponderação. Nos três procedimentos o escore de propensão é um componente fundamental, porém diferem na forma em que os pares são escolhidos e no cálculo do ATT.

Nesta pesquisa optamos por utilizar dois métodos de pareamento: i) método do vizinho mais próximo “*neighbor*” (para 3 vizinhos), que seleciona para cada elemento no grupo tratado, três indivíduos do grupo de controle, os mais próximos em termos

de escore de propensão, permitindo repetições, se necessário, em que cada observação do grupo de controle pode ser comparada mais de uma vez; ii) a regressão ajustada ponderada pelo IPWRA.

Este modelo é uma abordagem robusta utilizada na análise de dados para calcular as médias de resultados previstos em diferentes níveis de tratamento. Ele incorpora uma combinação de ponderação e regressão, sendo uma estratégia de estimação em dois estágios. Nesse método, os pesos são as probabilidades inversas estimadas de tratamento, obtidas através de técnicas de pareamento, neste caso o PSM. O IPWRA visa contornar o problema da má especificação dos modelos de tratamento e de resultado, garantindo estimativas consistentes mesmo quando apenas uma das equações é especificada corretamente. Essa consistência é alcançada por meio da utilização dos pesos do *Inverse Probability Weighted* (IPW) para estimar os coeficientes de regressão corrigidos, realizando posteriormente o ajuste de regressão com controles também no modelo de resultado. Assim, o IPWRA é considerado uma estratégia duplamente robusta, proporcionando maior confiabilidade aos resultados, especialmente em situações de especificação inadequada dos modelos. Ao incluir as mesmas variáveis na seleção e nos modelos de resultado, o IPWRA busca garantir a coerência e a sensibilidade na avaliação do efeito da variável de tratamento nas variáveis de resultado, destacando-se como uma ferramenta valiosa na análise de dados complexos e na tomada de decisões baseadas em evidências empíricas (MACÊDO, 2022; MASTRANGELO, MAIA, 2021; TAN, 2010; CALDEIRA, 2019; BANG, ROBINS, 2005; HIRANO, IMBENS, 2001; ROBINS, ROTNITZKY, 1995; ROBINS; ROTNITZKY; ZHAO, 1995; RUBIN, 1974).

A seção seguinte apresenta os resultados e discussões acerca dos modelos aplicados e, por fim, tem-se a conclusão da pesquisa.

4 RESULTADOS

Nesta seção são apresentadas as estimativas do perfil do consumidor *on-line* no Brasil e também o impacto da utilização de redes sociais na probabilidade de compra *on-line*. Optou-se pela apresentação desses resultados em subseções distintas, cada uma direcionada a uma das estratégias de estimação escolhidas (regressão preditiva – *Probit*, modelo PSM para análise causal e verificação de robustez – IPWRA).

4.1 *Probit*

Conforme descrito na seção anterior, a estratégia empírica deste estudo iniciou por estimar o modelo *Probit* cuja variável dependente é binária e indica se o indivíduo realizou compras *on-line* no último ano. O modelo é especificado considerando características pessoais, socioeconômicas, geográficas e comportamentais como variáveis independentes, bem como algumas interações entre essas características.

As interações fornecem boas informações acerca do perfil do consumidor e das diferenças entre gêneros. Realizamos interações entre gênero e *status* de atividade econômica (PEA), renda familiar, etnia, área urbana ou área rural, região de residência, utilização de redes sociais e transações financeiras realizadas pela internet. Ademais, para avaliar o efeito da idade na probabilidade de realizar compras *on-line*, considera-se um termo quadrático que permite capturar uma possível não linearidade nessa relação. Os resultados da estimativa são apresentados na Tabela 2.

Tabela 2 - Modelo Probit

	Variável dependente: comprou <i>on-line</i> no último ano			
	Coeficiente	Erro padrão	Efeito Marginal	Erro padrão
Ano				
2016	0,0403*	0,0199	0,0151*	0,0074
2017	-0,0793***	0,0194	-0,0290***	0,0071
2018	-0,0777***	0,0195	-0,0284***	0,0071
2019	0,0459*	0,0192	0,0172*	0,0072
2021	0,0920***	0,0199	0,0348***	0,0075
2022	-0,0470*	0,0205	-0,0173*	0,0075
Sexo masc.			0,0388***	0,0039
PEA				
Fora da força de trabalho			-0,0480***	0,0050
<i>Mulher</i>				
Fora da força de trabalho	-0,1230***	0,0163		
<i>Homem</i>				
Na força de trabalho	1.3491***	0,0712		
Fora da força de trabalho	1.2080***	0,0744		
Renda familiar				
Mais de 1 SM até 2 SM			0,0495***	0,0046
Mais de 2 SM até 3 SM			0,1026***	0,0056
Mais de 3 SM até 5 SM			0,1503***	0,0065
Mais de 5 SM			0,2170***	0,0083
<i>Mulher</i>				
Mais de 1 SM até 2 SM	0,1518***	0,0173		
Mais de 2 SM até 3 SM	0,2866***	0,0206		
Mais de 3 SM até 5 SM	0,3981***	0,0237		
Mais de 5 SM	0,5905***	0,0307		
<i>Homem</i>				
Até 1 SM	-0,5488***	0,0282		
Mais de 1 SM até 2 SM	-0,4231***	0,0278		
Mais de 2 SM até 3 SM	-0,2738***	0,0287		
Mais de 3 SM até 5 SM	-0,1380***	0,0300		
Mais de 5 SM	0 (omitida)			
Escolaridade				
Fundamental	0,1359***	0,0384	0,0399***	0,0107
Médio	0,5422***	0,0379	0,1814***	0,0107
Superior	0,9448***	0,0395	0,3401***	0,0115
Idade				
Idade	0,0206***	0,0023	-0,0007***	0,0001
Idade ²	-0,0003***	0,0000		
Etnia				
Preta			-0,0250***	0,0059
Parda			-0,0199***	0,0045
Amarela			0,0038	0,0120
Indígena			-0,0697***	0,0149
<i>Mulher</i>				
Preta	-0,0383	0,0223		
Parda	-0,0202	0,0167		
Amarela	-0,0299	0,0403		
Indígena	-0,1900**	0,0641		
<i>Homem</i>				

Branca	0,1947***	0,0543		
Preta	0,0928	0,0556		
Parda	0,1017	0,0535		
Amarela	0,2529***	0,0713		
Indígena	0 (omitida)			
Área				
Rural			-0,0624***	0,0063
<i>Mulher</i>				
Rural	-0,1821***	0,0248		
<i>Homem</i>				
Urbana	0,1622***	0,0266		
Rural	0 (omitida)			
Região				
Nordeste			-0,0906***	0,0051
Sul			-0,0493***	0,0062
Norte			-0,1474***	0,0059
Centro-Oeste			-0,0836***	0,0068
<i>Mulher</i>				
Nordeste	-0,2743***	0,0186		
Sul	-0,1371***	0,0220		
Norte	-0,4313***	0,0225		
Centro-Oeste	-0,2353***	0,0246		
<i>Homem</i>				
Sudeste	0,2022***	0,0265		
Nordeste	0,0049	0,0270		
Sul	0,0848**	0,0307		
Norte	-0,1652***	0,0295		
Centro-Oeste	0 (omitida)			
Redes Sociais				
Usou redes sociais			0,1179***	0,0047
<i>Mulher</i>				
Usou redes sociais	0,3629***	0,0203		
<i>Homem</i>				
Não usou redes sociais	-0,2942***	0,0186		
Usou redes sociais	0 (omitida)			
Transações financeiras				
Realizou transação financeira			0,3342***	0,0042
<i>Mulher</i>				
Realizou transação financeira	0,8776***	0,0155		
<i>Homem</i>				
Não realizou transação financeira	-0,8970***	0,0161		
Realizou transação financeira	0 (omitida)			
<i>McFadden's R²</i>	0,230	Wald χ^2	20699,28	0,0000
Observações	82,339			

Fonte: elaboração própria.

Nota: o *Software Stata*® gera uma categoria de referência para cada variável categórica, ver Tabela 6 no apêndice A. As entradas são coeficientes estimados por máxima verossimilhança. Erros padrão robustos para heterocedasticidade. Para a significância estatística: (*) < 10%, (**) < 5% e (***) < 1%.

A análise dos resultados é feita predominantemente por meio de probabilidades estimadas e efeitos marginais, como sugerido por Long e Freese (2014), ainda, é importante ressaltar que, devido à natureza do modelo *Probit*, não é possível estabelecer relações causais diretas entre as características analisadas e a probabilidade de compra *on-line*. No entanto, as análises fornecem previsões e uma compreensão mais profunda das associações entre essas características e o comportamento de compra *on-line*.

Antes de realizar essa análise, é necessário fazer alguns comentários breves sobre o ajuste do modelo. A taxa de classificação de casos que foram corretamente classificados pelo modelo, considerando tanto os verdadeiros positivos quanto os verdadeiros negativos é de 75,78%. Dentre essa taxa, 70,15% dos casos classificados como positivos pelo modelo realmente corresponderam a indivíduos que realizaram compras *on-line*. E, para os casos classificados como negativos pelo modelo, 78,18% realmente corresponderam a indivíduos que não realizaram compras *on-line*. O teste de razão de verossimilhança (*Likelihood Ratio Test* – LR), sugere que há evidências estatisticamente significativas para rejeitar a hipótese nula de que o modelo restrito (com apenas a constante) é adequado para prever a probabilidade de compra *on-line*. Isso significa que o modelo completo, que inclui todas as variáveis independentes, é estatisticamente superior ao modelo com apenas a constante, indicando que as variáveis independentes têm um impacto significativo na previsão da variável dependente. Já o R^2 de *McFadden* apresentou valor de 0,23, que indica um ajuste próximo ao considerado por *McFadden* (1974) como "muito bom", que está entre 0,2 e 0,4.

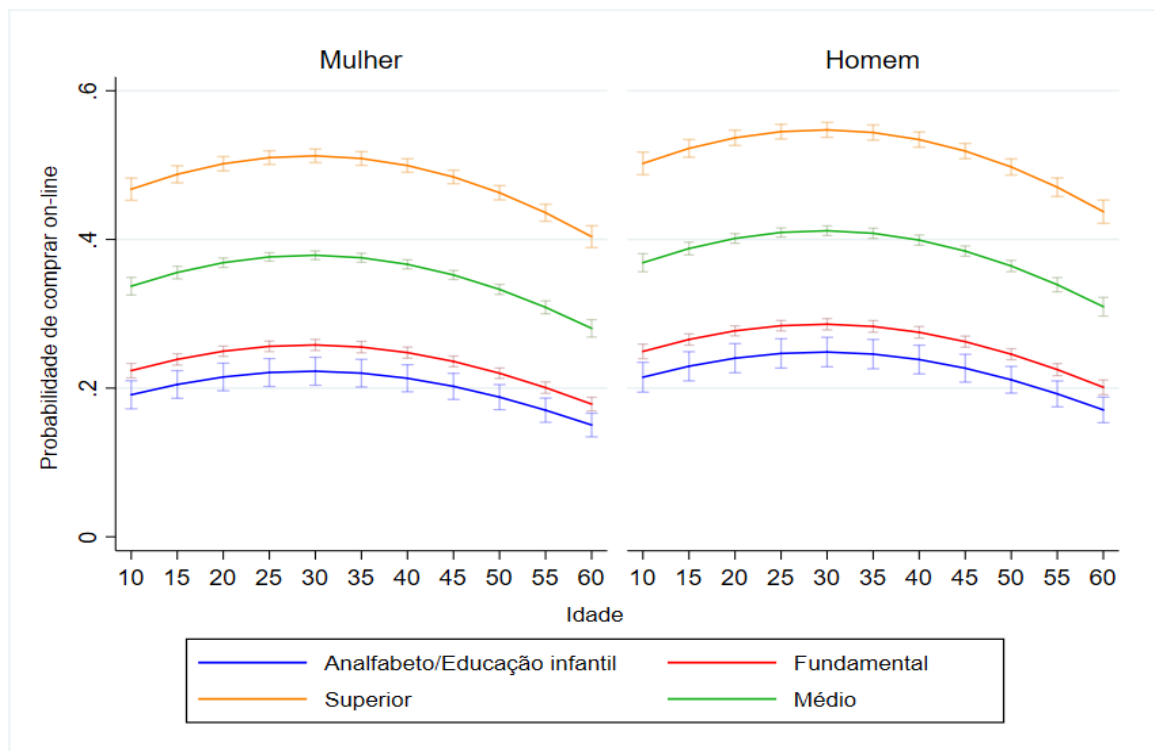
Inicia-se a análise dos resultados pelas *dummies* de “ano” que indicam um aumento na probabilidade de consumo *on-line* em 2021, provavelmente impulsionado pelos efeitos da pandemia que aceleraram o crescimento do *e-commerce* devido a necessidade do isolamento social. Em 2022, há um declínio, que pode ser justificado pelo retorno das atividades presenciais. Ainda, na Tabela 2, analisando a variável de gênero, ser do sexo masculino indica uma probabilidade maior de realizar compras *on-line* de 3,8 pontos percentuais, quando comparado ao sexo feminino. Apesar de indicar uma pequena diferença, essa relação de gênero também se verifica quando analisamos as interações com as demais variáveis independentes.

Para a análise do efeito da idade na probabilidade de realizar compras pela internet, utilizamos duas variáveis: idade e idade². A primeira variável apresenta sinal

positivo e o termo quadrático sinal negativo. Assim, capturamos os efeitos na margem da idade, ou seja, a probabilidade de consumo *on-line* aumenta com a idade, mas a taxas decrescentes. Cabe destacar que, conforme já abordado, a idade é um fator determinante para o comércio eletrônico, entretanto há uma tendência para que o ponto de ruptura entre consumidores e não consumidores (*on-line*) seja maior, ou seja, indivíduos com faixas etárias maiores nas próximas décadas serão mais ativos comparativamente aos indivíduos da mesma faixa etária atualmente (RYBACZEWSKA; SPARKS, 2022). Nos gráficos a seguir é possível visualizar essa relação.

Alguns estudos apontam que o usuário médio da internet pertence às camadas mais privilegiadas da população, com renda familiar bem acima da média, e o nível educacional tendendo para ensino superior (MORGADO, 2003). Aqui, da mesma forma, observamos que um maior nível de escolaridade dos indivíduos aumenta a probabilidade de consumo *on-line* (ver Gráfico 1).

Gráfico 1 - Probabilidade de realizar compras *on-line* de acordo com o nível de escolaridade para cada gênero no Brasil



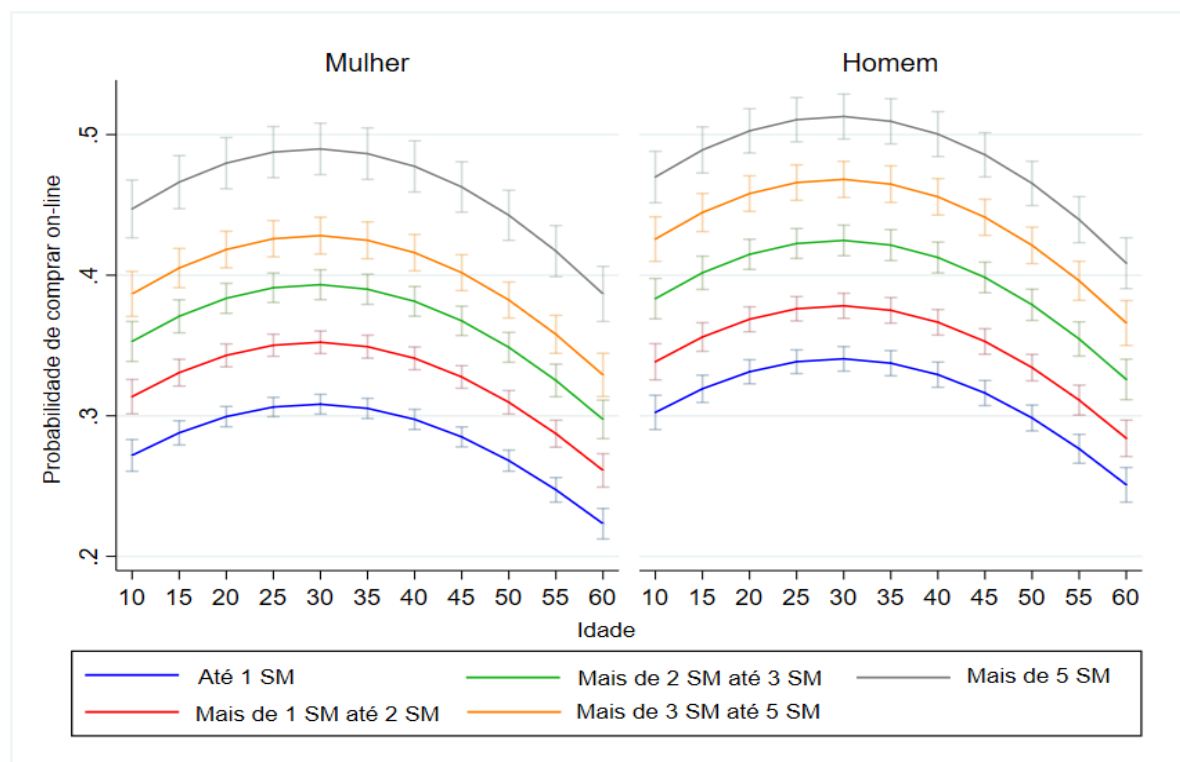
Fonte: gráfico gerado através do *Software Stata*®.

Nota: as entradas são probabilidades previstas estimadas pela máxima verossimilhança num modelo *Probit*. Os intervalos foram calculados usando erros padrão robustos, 95% de confiança. A idade é medida em anos.

Cabe destacar que, na Tabela 2, a relação entre escolaridade e a probabilidade de consumo *on-line*, utiliza “Analfabeto/Educação infantil” como a categoria de referência, e todos os coeficientes positivos da Tabela, para as categorias de escolaridade, indicam que, à medida que o nível de escolaridade aumenta (do Analfabeto/Educação infantil ao Superior), a probabilidade estimada de realizar compras *on-line* também aumenta. Os efeitos marginais indicam a proporção desse efeito, por exemplo, um indivíduo com ensino fundamental tem maior probabilidade de consumo em 3,99 pontos percentuais, quando comparado com a categoria de referência. A categoria de ensino “Superior” apresenta o maior efeito marginal, 34 pontos percentuais.

A escolaridade está correlacionada com maiores níveis de renda e, por conseguinte, com maiores níveis de consumo. Utilizamos a renda familiar e idade a partir de 10 anos a fim de capturar uma possível participação dos indivíduos de menoridade no consumo *on-line* das famílias, conforme se verifica no Gráfico 2.

Gráfico 2 - Probabilidade de realizar compras *on-line* de acordo com o nível de renda familiar para cada gênero e idade



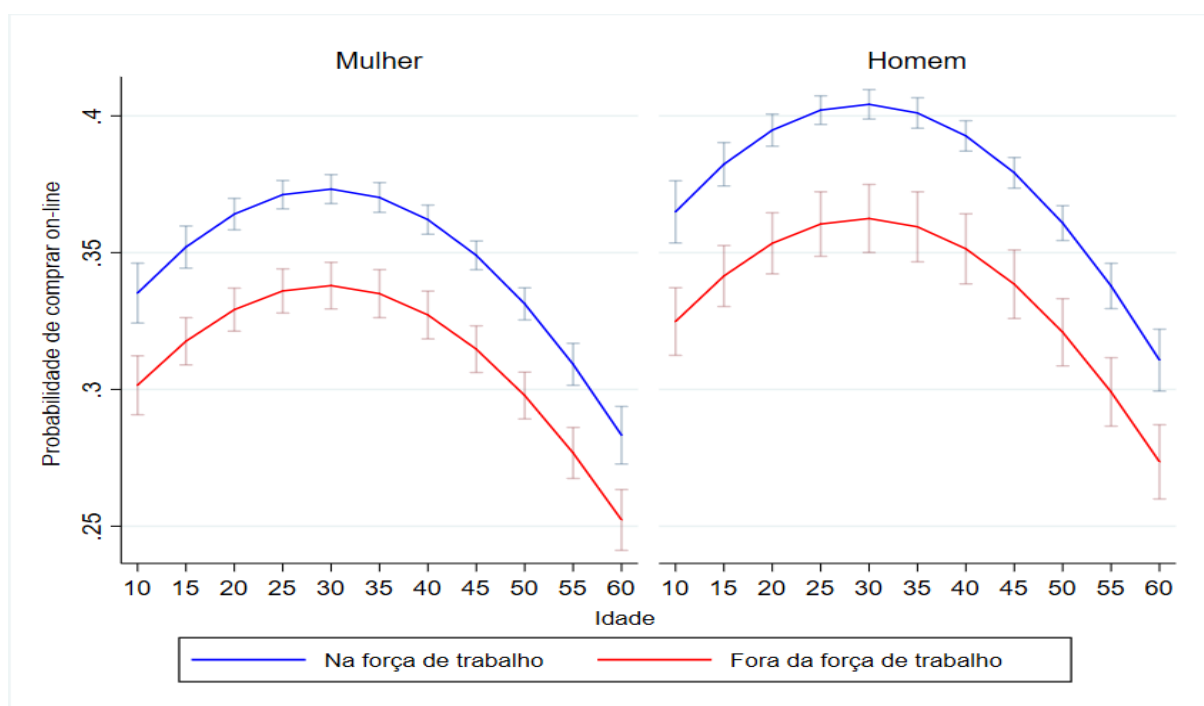
Fonte: gráfico gerado através do *Software Stata*®.

Nota: as entradas são probabilidades previstas estimadas pela máxima verossimilhança num modelo *Probit*. Os intervalos foram calculados usando erros padrão robustos, 95% de confiança. A idade é medida em anos.

Neste caso a renda familiar está dividida em cinco faixas baseadas no salário mínimo vigente de cada ano da coleta de dados. Analisando a Tabela 2, temos duas categorias de referência: (i) para as mulheres: receber até 1 salário mínimo; (ii) para os homens: receber mais de 5 salários mínimos. Os resultados da tabela indicam, assim como outras pesquisas (SONG; SUN, 2020; GALINARI et al., 2015; LIMA FILHO et al., 2012; MORGADO, 2003; LI; KUO; RUSSELL, 1999) que o nível de renda tem uma relação positiva com a propensão de consumo, quanto maior a renda familiar maior a probabilidade de consumo *on-line*.

A variável PEA, é uma *dummy* que indica se o indivíduo está na força de trabalho, ou não. Seguindo a lógica da renda familiar, indivíduos que estão na força de trabalho têm mais condições financeiras para consumir *on-line* quando comparados a indivíduos que estão fora do mercado de trabalho (ver Gráfico 3).

Gráfico 3 - Probabilidade de realizar compras *on-line* de acordo com o gênero, idade e *status* de atividade econômica - PEA



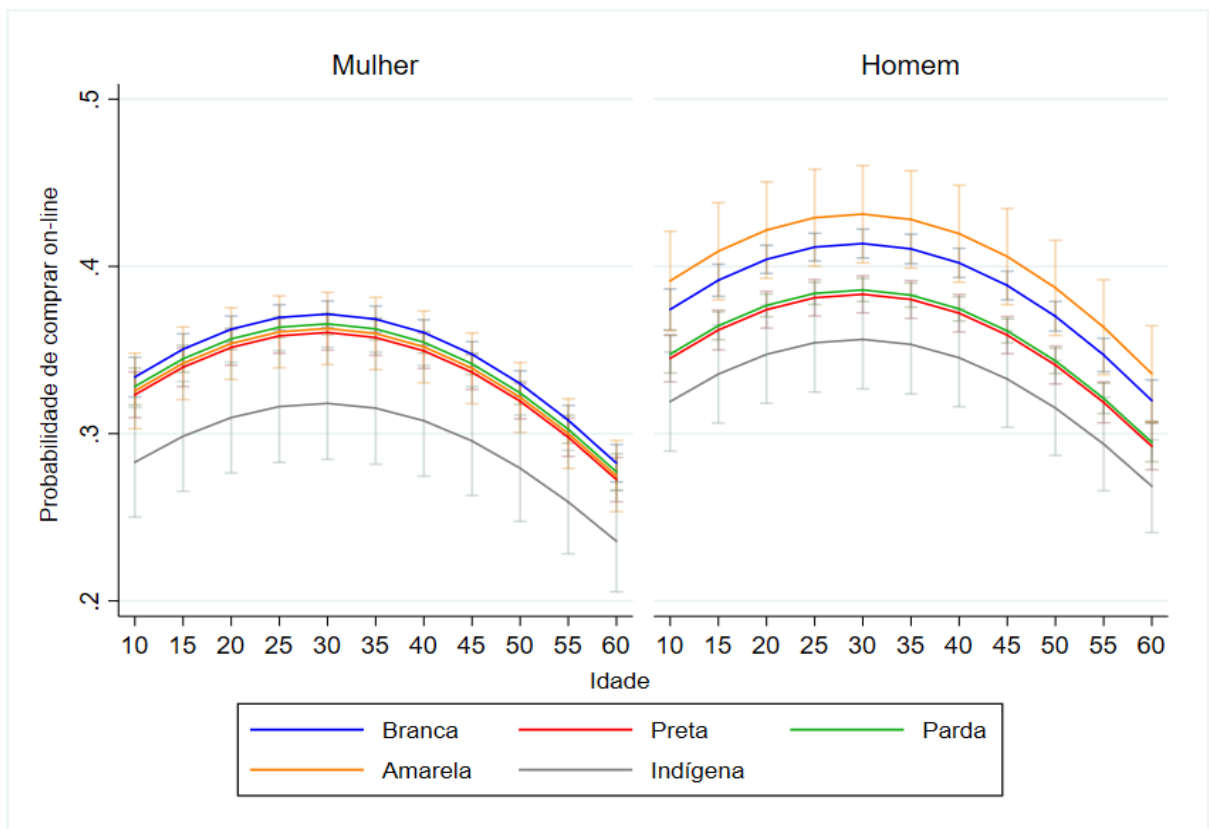
Fonte: gráfico gerado através do *Software Stata*®.

Nota: as entradas são probabilidades previstas estimadas pela máxima verossimilhança num modelo *Probit*. Os intervalos foram calculados usando erros padrão robustos, 95% de confiança. A idade é medida em anos.

Na Tabela 2, a categoria de referência é “Mulher” e “Na força de trabalho”, de acordo com os dados estimados pelo modelo, podemos concluir que os indivíduos que estão na força de trabalho tendem a ser mais propensos a fazer compras *on-line*, para ambos os sexos.

Ainda, acerca da análise entre a etnia dos indivíduos e a probabilidade de consumo *on-line*, podemos inferir que, a etnia pode estar associada a diferenças na probabilidade de consumo. No entanto, é importante notar que os efeitos são relativamente pequenos e para algumas categorias os resultados não são estatisticamente significativos (ver Gráfico 4).

Gráfico 4 - Probabilidade de realizar compras *on-line* de acordo com o gênero, idade e etnia, no Brasil



Fonte: gráfico gerado através do *Software Stata*®.

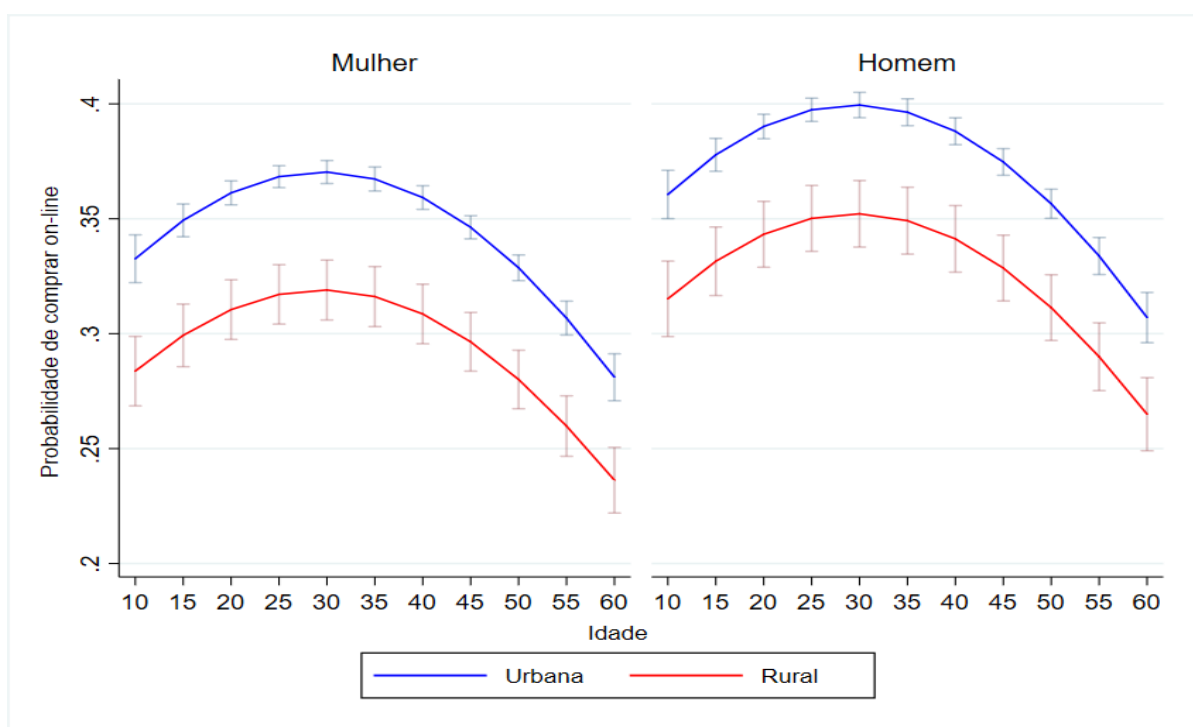
Nota: as entradas são probabilidades previstas estimadas pela máxima verossimilhança num modelo *Probit*. Os intervalos foram calculados usando erros padrão robustos, 95% de confiança. A idade é medida em anos.

Para o grupo das mulheres a categoria de referência é a “Branca”, e todas as demais categorias estão com sinal negativo (ver Tabela 2), indicando que a maior probabilidade de comprar *on-line* entre as mulheres está na categoria de etnia

“Branca”. Já no grupo dos homens, a categoria de referência é “Indígena”, e todas as demais categorias apresentam sinal positivo, indicando que possuem maior propensão de compra *on-line* quando comparado com a etnia “Indígena”.

Os fatores geográficos são analisados através de duas variáveis: regiões brasileiras e área de residência (urbana ou rural). Residir na área urbana afeta positivamente a probabilidade de realizar compras *on-line* para ambos os sexos, isso pode ter relação com o isolamento da área rural às TIC. Esse resultado está de acordo com algumas pesquisas (GALINARI et al., 2015; IPEA, 2011). O Gráfico 5 ilustra esse resultado entre os gêneros.

Gráfico 5 - Probabilidade de realizar compras *on-line* de acordo com o gênero, idade e a área de residência



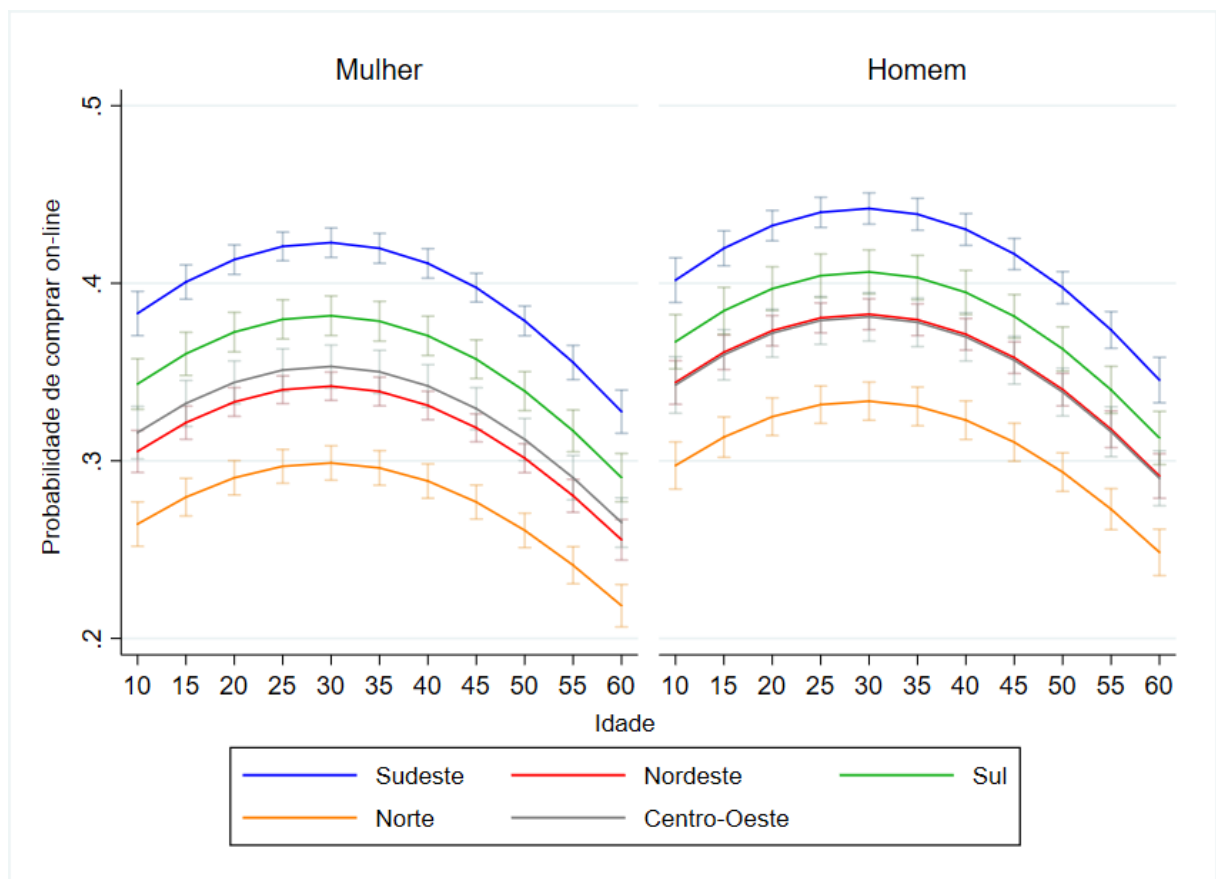
Fonte: gráfico gerado através do *Software Stata*®.

Nota: as entradas são probabilidades previstas estimadas pela máxima verossimilhança num modelo *Probit*. Os intervalos foram calculados usando erros padrão robustos, 95% de confiança. A idade é medida em anos.

Além da área, a região de residência pode indicar diferenças na probabilidade de compra *on-line*, conforme segue no Gráfico 6. Nesse aspecto, quando analisamos as regiões geográficas do Brasil, ambos os sexos apresentam tendências semelhantes. A região com maior probabilidade de consumo *on-line* é o Sudeste, seguido pelo Sul. O Norte apresenta menor probabilidade comparado às demais

regiões. Por fim, o Nordeste e Centro-Oeste apresentam resultados muito próximos. O Sudeste representa pouco mais de 40% da população brasileira, e conta com os 3 estados mais populosos do país, segundo o Instituto Brasileiro de Geografia e Estatística (IBGE, 2023), portanto esse resultado está conforme esperado.

Gráfico 6 - Probabilidade de realizar compras *on-line* de acordo com o gênero, idade e região geográfica do Brasil



Fonte: gráfico gerado através do *Software Stata®*.

Nota: as entradas são probabilidades previstas estimadas pela máxima verossimilhança num modelo *Probit*. Os intervalos foram calculados usando erros padrão robustos, 95% de confiança. A idade é medida em anos.

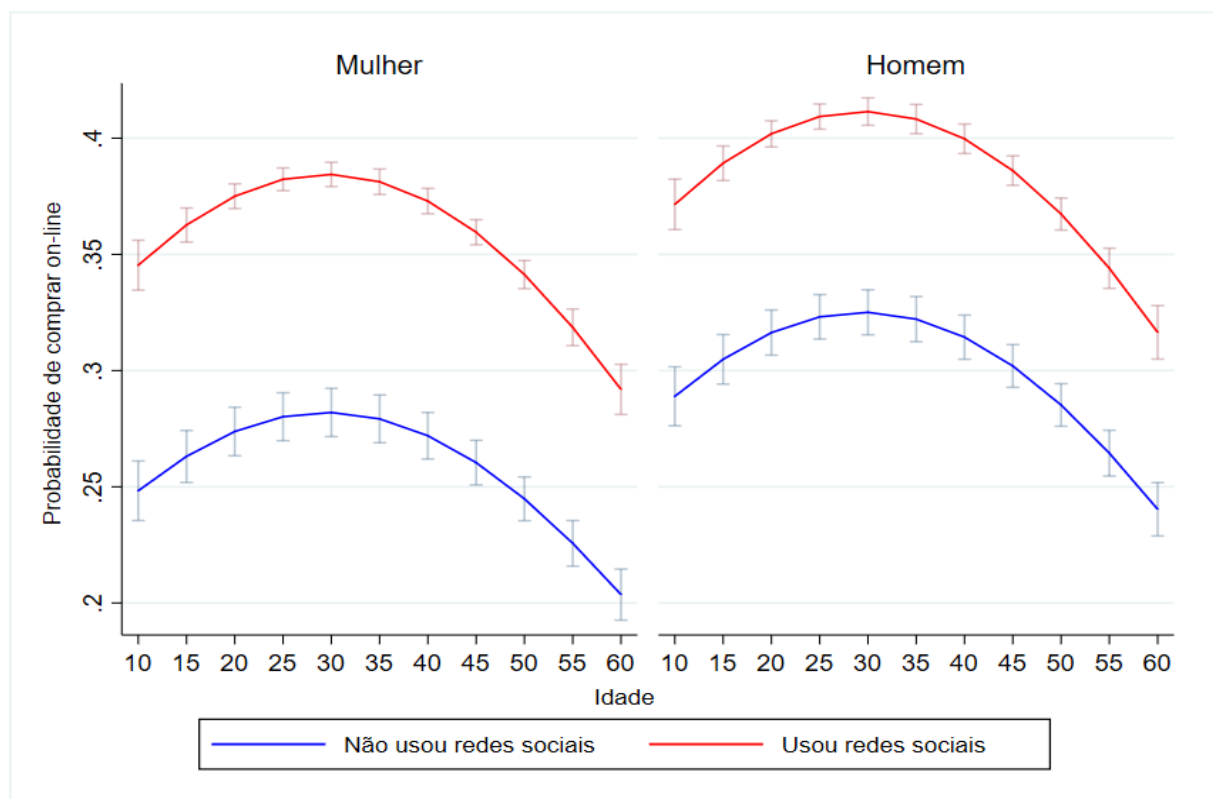
Na Tabela 2, os efeitos marginais são todos negativos, pois a categoria de comparação é o “Sudeste”. Para a categoria “Norte”, o efeito marginal é de –14,74 pontos percentuais. Isso sugere que os indivíduos que residem nesta região têm uma probabilidade menor de realizar compras *on-line* em comparação com os indivíduos do Sudeste, mantendo todas as outras variáveis constantes. A Tabela 2 apresenta resultado análogo para as demais regiões do país. Esse resultado vai de encontro a outros estudos realizados no Brasil, a exemplo de Cardoso, Kawamoto e Massuda

(2019), em que se verificou uma concentração das empresas de comércio eletrônico na região Sudeste.

Características relacionadas ao comportamento dos indivíduos perante a internet também podem contribuir para o crescimento ou declínio do comércio eletrônico. Neste modelo, temos duas variáveis independentes para prever a probabilidade de consumo *on-line* com base na sua utilização da internet: (i) utilização de redes sociais e (ii) realização de transações financeiras pela internet.

As redes sociais alteraram a forma como os indivíduos se comunicam e compartilham informações. Atualmente as redes sociais, através dos *digitais influencers*, *streamers*, propagandas e *lives* facilitam o acesso ao comércio eletrônico (OPINIONBOX, 2023; ANDRADE; SILVA, 2017; YADAV; RAHMAN, 2017; GALINARI et al., 2015; LINDA, 2010). Essas novas formas de comunicação também impactam e influenciam a forma como o consumidor se comporta, ver Gráfico 7.

Gráfico 7 - Probabilidade de realizar compras *on-line* de acordo com o gênero, idade e utilização de redes sociais



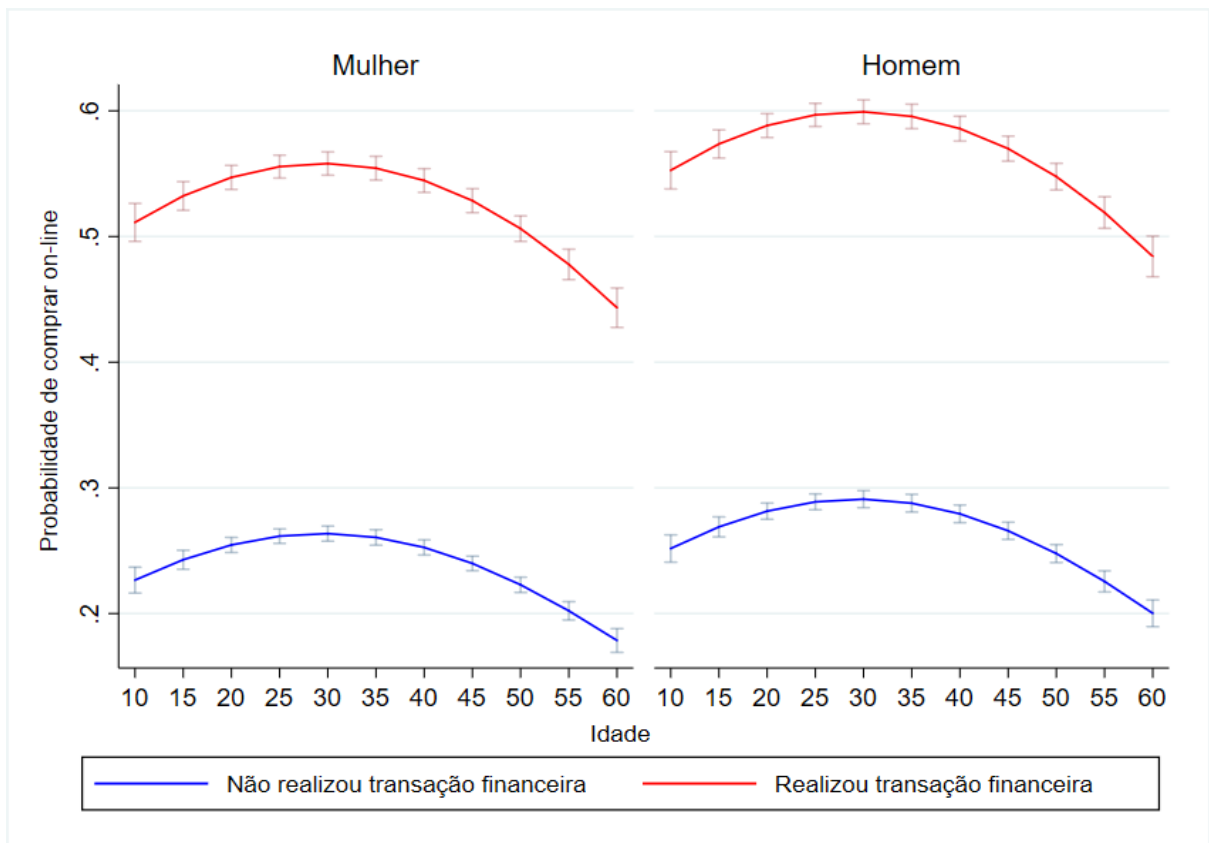
Fonte: gráfico gerado através do *Software Stata*®.

Nota: as entradas são probabilidades previstas estimadas pela máxima verossimilhança num modelo *Probit*. Os intervalos foram calculados usando erros padrão robustos, 95% de confiança. A idade é medida em anos.

Na Tabela 2, podemos analisar os efeitos marginais para a categoria "Usou redes sociais". Os indivíduos que utilizam redes sociais têm maior probabilidade de consumo *on-line*, esse impacto é de 11,79 pontos percentuais em comparação com os indivíduos que não utilizam redes sociais. Essa relação será detalhada na próxima seção, a fim de verificar se existe uma relação causal entre o uso das redes sociais e o comércio eletrônico.

Assim como as redes sociais, a bancarização é outro fator relevante quando tratamos de *e-commerce* (MORGADO, 2003). Realizar pagamentos e fazer transações financeiras pela internet têm relação positiva e significativa com a propensão a consumir *on-line*.

Gráfico 8 - Probabilidade de realizar compras *on-line* de acordo com o gênero, idade e se realizou transações financeiras pela internet



Fonte: gráfico gerado através do *Software Stata*®.

Nota: as entradas são probabilidades previstas estimadas pela máxima verossimilhança num modelo *Probit*. Os intervalos foram calculados usando erros padrão robustos, 95% de confiança. A idade é medida em anos.

Na Tabela 2 há duas categorias de referência: para o grupo das mulheres “Não realizou transação financeira”, portanto, quando analisamos o coeficiente estimado para a categoria que realiza transações financeiras pela internet o sinal é positivo, indicando uma maior propensão ao consumo *on-line*; para o grupo dos homens a categoria de referência é “Usou redes sociais”, de forma análoga, o sinal é negativo, indicando que quem não realiza transações financeiras pela internet tem menor probabilidade ao consumo *on-line*.

Para concluir, vale salientar que, o efeito marginal de realizar transações financeiras pela internet é de 33,42 pontos percentuais, indicando como essa variável pode influenciar de forma positiva e significativa na propensão de compra dos indivíduos. O efeito de realizar transações financeira pela internet possui um impacto significativo na decisão de realizar compras *on-line*, esse resultado está em conformidade com outro estudo empírico que também evidenciou essa relação (ALSHWEESH; BANDI, 2022).

4.2 Propensity Score Matching (PSM)

A metodologia do PSM nos possibilita avançar na interpretação dos resultados previamente obtidos pelo *Probit*, pois busca investigar se existe uma relação causal entre o uso de redes sociais e a probabilidade de indivíduos realizarem compras pela internet. Para estimar o modelo utiliza-se o *Software Stata*®.

Conforme já apresentado na seção de metodologia, o PSM começa com o modelo *logit* para cálculo do escore de propensão, a Tabela 7, no Apêndice B, contém todos os resultados da estimação. Cabe destacar que quase a totalidade dos indivíduos estão no suporte comum, foram excluídas apenas 76 pessoas. No total têm-se 69.295 observações, conforme Tabela 3¹¹.

Tabela 3 - PSM - Suporte comum

Tratamento: Redes Sociais	Fora do suporte	No suporte	Total
Controle	5	13.587	13.592
Tratado	71	55.632	55.703
Total	76	69.219	69.295

Fonte: adaptação da tabela extraída do *Software Stata*®, utilizando o comando *psmatch2*.

¹¹ Ver Gráfico 9 – Suporte Comum, no Apêndice B.

Ainda, apresenta-se uma análise detalhada dos indicadores de ajuste e testes acerca da aplicação do modelo PSM¹². Os resultados são exibidos na Tabela 4, que resume os principais indicadores de interesse para este estudo.

Tabela 4 - PSM - *Pstest*

Amostra	Pseudo R2	LR Chi ²	P>Chi ²	Viés médio	R
Não pareado	0,120	8212,43	0,000	13,8	0,90
Pareado	0,001	77,35	0,000	0,7	0,99

Fonte: adaptação da tabela extraída do *Software Stata*®, utilizando o comando *pstest*.

A análise dos indicadores na Tabela 4 revela informações cruciais sobre a eficácia do PSM na redução do viés de seleção e no ajuste das amostras. Ao comparar os resultados antes e depois do pareamento, podemos observar mudanças significativas em diferentes métricas. Primeiramente, o indicador R, que representa a razão da variância entre os grupos após o PSM em relação à variância antes do PSM, mostrou uma melhoria substancial, passando de 0,90 para 0,99. Isso indica que o PSM foi altamente eficaz em reduzir a variabilidade entre os grupos, contribuindo para uma comparação mais precisa e confiável.

O valor do teste qui-quadrado de *Pearson* (LR chi²) também diminuiu consideravelmente após o pareamento, indicando uma redução significativa na disparidade entre as distribuições das variáveis nos grupos. Esse resultado, junto com o valor “p” associado ao teste qui-quadrado (p>chi²), que demonstrou uma diferença estatisticamente significativa entre os grupos antes e depois do PSM, reforça a eficácia da técnica na equalização das amostras. Outro aspecto relevante é o viés médio, que representa a diferença média entre os grupos antes e depois do PSM para a variável de interesse. A redução substancial do viés médio de 13,8 para 0,7 indica que o PSM foi capaz de minimizar significativamente o viés de seleção, aumentando a confiabilidade das comparações realizadas no estudo.

O resultado principal do modelo é apresentado na Tabela 5. O cálculo do efeito médio do tratamento sobre os tratados é estimado de duas formas, através do PSM (ATT) e do IPWRA (*Average Treatment Effect on the Treated* - ATET). Podemos observar que os valores encontrados são muito semelhantes, respectivamente 13,42 p.p. e 13,19 p.p., indicando um bom ajuste do modelo e tornando os resultados aqui

¹² Ver Gráfico 10 – PSM *Matching*, no Apêndice B.

expostos mais robustos. Portanto podemos inferir que a probabilidade de compra *on-line* aumenta em 43,5% com a utilização das redes sociais.

Tabela 5 - Efeito do tratamento no grupo tratado sob diferentes métodos

	Variável dependente: fez compras <i>on-line</i> no último ano.					
	Tratado	Controle	Diferença	Estatística T	Erro Padrão	P > z
Não pareado	0,4433	0,2070	0,2362	51,42	-	
PSM ATT	0,4428	0,3085	0,1342	22,69	-	0,000
IPWRA <i>POmean</i>	0,3864	0,2578	0,1286	-	-	0,000
IPWRA ATET	-	-	0,1319	-	0,0049	0,000

Fonte: adaptação da tabela extraída do *Software Stata*®.

Observação da tabela: *POmean*: indica a média ponderada da variável de interesse com base nos escores de propensão.

Da mesma forma, pelo modelo IPWRA, observamos que os indivíduos que não utilizam redes sociais (controle) têm uma probabilidade de compra *on-line* de 25,78 p.p., entretanto utilizar rede social aumenta a probabilidade de realizar compras *on-line* em mais 13,19 p.p. Ou seja, a probabilidade de um indivíduo comprar *on-line*, dado que utiliza rede social, aumenta em 51%.

Em resumo, os resultados sugerem que o modelo PSM foi eficaz em reduzir o viés de seleção entre os grupos, como indicado pela redução significativa no viés médio, o baixo valor de LR χ^2 e $p > \chi^2$ também indica que houve uma melhoria estatisticamente significativa na comparação entre os grupos antes e depois do PSM. Os resultados apresentados com o modelo IPWRA também indicam a robustez do modelo estimado.

Portanto, podemos inferir que as redes sociais desempenham um papel fundamental no comércio eletrônico. Há diversos fatores que podem estar relacionados ao uso das redes sociais e o aumento da probabilidade de compra *on-line*, tais como: i) os usuários das redes sociais passam mais tempo no celular, ou seja, estão mais expostos às propagandas *on-line* (de forma direta ou através dos *influencers*); ii) maior confiança em suas habilidades de navegação nas redes e na internet em geral, conseqüentemente possuem menos receio em usar a internet para compras e, iii) apresentam maior confiança para realizar pagamentos *on-line*, tornando o processo de compra mais fácil e acessível. Os resultados aqui encontrados estão de acordo com outros estudos empíricos (SAHA; WANG; OH, 2023; SAHNEY, 2021; SHI; WANG, 2020; ZHANG et al., 2017; MADNI, 2015; HAJLI, 2014).

Essa pesquisa se destaca por trazer avanços relativos a pesquisas anteriores, a exemplo da pesquisa IPEA (2011), que descreve o perfil do consumidor *on-line* brasileiro, convergindo com as descobertas presentes neste trabalho. Contudo, há uma diferença importante entre ambas, a relação entre o uso das redes sociais e o comércio eletrônico. A referida pesquisa do IPEA (2011) não evidencia significância estatística entre o uso de redes sociais e o consumo *on-line*. Portanto a presente pesquisa avança ao examinar essa relação, adotando uma metodologia diferente e dados mais atualizados, possibilitando uma análise mais robusta da relação entre o uso de redes sociais e o comércio eletrônico. Além de definir o perfil do consumidor *on-line* brasileiro atualizado.

5 CONCLUSÕES

Durante os anos analisados, o consumo *on-line* apresentou crescimento contínuo, com exceção do ano de 2022 em que houve um pequeno declínio (CGI.br, 2023). Isso reforça como o comércio eletrônico mudou os hábitos tradicionais de compra e se consolidou com uma nova forma de consumo. Conforme proposto, este estudo buscou analisar como as características dos indivíduos afetam a probabilidade de realizar compras *on-line*, e assim, por se tratar de um modelo de previsão, os resultados do *Probit*, trazem informações relevantes para tentar prever um perfil do consumidor, com as características incluídas no modelo, que tenha a maior probabilidade de realizar compras *on-line*, no contexto brasileiro.

As características incluídas no modelo, como variáveis independentes, em sua maioria, apresentaram resultados estatisticamente significativos, com exceção das categorias “etnia” e “região geográfica”. As características que apresentaram os maiores efeitos marginais positivos e coeficientes estimados significativos - consequentemente, formam um perfil com maior propensão de consumo *on-line* - são: sexo=1 (homem); estar trabalhando; utilizar redes sociais; realizar transações financeiras na internet; residir em área urbana e na região sudeste; ter renda superior a 5 salários mínimos e nível de escolaridade “Superior”.

Vale destacar alguns aspectos acerca deste perfil mencionado. Ser do sexo masculino apresenta um efeito maior na propensão do consumo, entretanto esse efeito marginal é pequeno quando comparado ao percentual total de indivíduos que compraram *on-line*, pela Tabela 1 o percentual é de 36,27% e o efeito de ser do sexo masculino aumenta a probabilidade de compra em 3,88 pontos percentuais. Portanto, essa diferença entre os gêneros talvez não seja a característica mais marcante na propensão do consumo *on-line*.

Em contrapartida, quando analisamos o efeito de utilizar redes sociais e realizar transações financeiras pela internet, por exemplo, esse impacto fica bem mais evidente, já que essas variáveis aumentam a probabilidade de compra em, respectivamente, 11,79 e 33,42 pontos percentuais. Dessa forma, as características que mais afetam a propensão a participar do comércio eletrônico são, por ordem decrescente: escolaridade nível “Superior” (34,01 p.p); transações financeiras na internet (33,42 p.p.); renda familiar “Mais de 5 SM” (21,70 p.p), e assim segue.

Ao empregar o modelo *Probit*, não é possível estabelecer uma relação causal direta entre as variáveis independentes do modelo e a decisão de realizar compras *on-line*. Portanto, recorreremos ao PSM como uma abordagem mais apropriada para fazer inferência causal. Os resultados obtidos por meio do PSM revelam uma relação positiva e significativa entre o uso das redes sociais e a probabilidade de realizar compras *on-line*. Especificamente, observamos que o impacto do uso das redes sociais aumenta a probabilidade de compra em 43,5%. Além disso, o modelo demonstra um bom ajuste e um teste confiável após o pareamento. Para reforçar a robustez dos resultados, realizamos uma análise de sensibilidade com o IPWRA, que confirma a consistência do modelo e aponta um impacto ainda maior das redes sociais, elevando a probabilidade de compra *on-line* para 51%, quando um indivíduo a utiliza ativamente. Esses resultados consolidam a importância do uso estratégico das redes sociais no contexto do comércio eletrônico, destacando seu papel significativo na decisão de compra dos consumidores.

Ademais, vale ressaltar que este estudo possui limitações. O modelo inicial de previsão (*probit*) apresenta um bom ajuste, entretanto há outros modelos econométricos que, por exemplo, utilizam *machine learning* para realizar previsões mais assertivas (ZHANG; YANG, 2022). Com vista do modelo PSM as limitações estão relacionadas a falta de dados para trabalhar com dados em painel, que permitiria fazer uma análise dinâmica e mais robusta acerca do tema, além de reduzir o viés de omissão de variáveis relevantes.

Por fim, essa pesquisa oferece uma base teórica e analítica para um futuro estudo que empregue um modelo econométrico mais robusto, utilizando dados em painel. Isso permitirá uma compreensão mais aprofundada das influências das características individuais na decisão de participar do comércio eletrônico no Brasil. Sugestões para trabalhos futuros incluem a expansão do estudo para o comércio exterior e mundial, a exemplo da pesquisa realizada em Taiwan que abrange o comércio nacional e transfronteiriço (HUANG et al., 2021), bem como a análise de questões regulatórias, elementos cruciais para o avanço do *e-commerce* brasileiro, a exemplo de Magalhães, Ferreira e Silva (2022).

REFERÊNCIAS

- AGUIAR, A. Do tradicional ao composável: a nova era do comércio eletrônico. **e-commercebrasil**, [s.l.], 17 nov. 2023. Disponível em: <https://www.e-commercebrasil.com.br/artigos/do-tradicional-ao-composavel-a-nova-era-do-comercio-eletronico>. Acesso em: 02 jul. 2024.
- ALSHWEESH, R.; BANDI, S. The Impact of E-Commerce on Consumer Purchasing Behavior: The Mediating Role of Financial Technology. **International Journal of Research and Review**, v. 9, n.2, p.479-499, 2022.
- ANDRADE, M. C. F. de; SILVA, N. G. O comércio eletrônico (e-commerce): um estudo com consumidores. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 7, n. 1, p. 98-111, jan./jun. 2017.
- ARCOVERDE, D. O que é ser um streamer, como é essa profissão e quais são os maiores do Brasil. **Netshow.me**, São Paulo, 17 mar. 2023. Disponível em: <https://netshow.me/blog/streamer/#:~:text=Um%20streamer%20%C3%A9%20um%20profissional,no%20Brasil%20e%20no%20mundo>. Acesso em: 15 nov. 2023.
- BALESTRIN, V. A.; DARIGO, E. B. *Marketing* digital como ferramenta de persuasão: estudo de caso e-commerce Magazine Luiza. In: XX SEMINÁRIO DE INICIAÇÃO CIENTÍFICA. VII SEMINÁRIO INTEGRADO DE ENSINO, PESQUISA E EXTENSÃO. V MOSTRA UNIVERSITÁRIA – SÃO MIGUEL DO OESTE, SC: UNOESC, 2014, P. 141. **Anais...** XX Seminário de Iniciação Científica. VII Seminário Integrado de Ensino, Pesquisa e Extensão. V Mostra universitária, 2014 Disponível em: <https://periodicos.unoesc.edu.br/siepe/article/view/5464/2817>. Acesso em: 31 mar. 2024.
- BANG, H.; ROBINS, J. M. Doubly robust estimation in missing data and causal inference models. **Biometrics**, v. 61, n. 4, p.962-973, dez. 2005.
- BARBOSA, R. M. Comércio eletrônico + redes sociais = comércio social. **NIC.br**, [s.l.], 11 jan. 2010. Disponível em: <https://nic.br/noticia/na-midia/comercio-eletronico-redes-sociais-comercio-social/>. Acesso em: 02 jul. 2024.
- BAYMARD INSTITUTE. 49 Cart Abandonment Rate Statistics 2023. 70.19% – average documented online shopping cart abandonment rate. **Baymard Institute**, Copenhagen, 11 jul. 2023. Disponível em: <https://baymard.com/lists/cart-abandonment-rate>. Acesso em: 05 nov. 2023.
- CALDERA, S. **Techniques for robustness and threats to inference**: inverse probability weighted regression adjustment. 2019. Disponível em: <http://sites.utexas.edu/prc/files/IPWRA.pdf>. Acesso em: 31 mar. 2024.
- CARDOSO, S.; KAWAMOTO, M. H.; MASSUDA, E. M. Comércio eletrônico: o varejo virtual brasileiro. **Revista Cesumar**, v. 24, n. 1, p. 117-134, jan./jun. 2019.
- CGI.br - COMITÊ GESTOR DA INTERNET NO BRASIL. **Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros: TIC Domicílios 2022**. 1ª ed. São Paulo: CGI.Br, 2023.

CUNHA, G. Comportamento do consumidor: como mapear e analisar os hábitos? **ESPM**, [s.l.], 22 abr. 2021. Disponível em: https://blog.espm.br/pos/comportamento-do-consumidor-como-mapear-e-analisar-os-habitos?utm_campaign=PMAX-RMKT_INSTITUCIONAL-POS-GRADUACAO-POA-2024_2&utm_source=adwords&utm_medium=ppc&utm_term=&hsa_kw=&hsa_net=adwords&hsa_grp=&hsa_cam=21380438341&hsa_acc=1176814211&hsa_tgt=&hsa_ver=3&hsa_ad=&hsa_mt=&hsa_src=x&gad_source=1&gclid=CjwKCAjwhlS0BhBqEiwADAUhcwSWaXwL5LjhpMkAltOXkGkY3Qe-XoTIRiBvhqflLBNkOcc4OEPjbxoCwk0QAvD_BwE. Acesso em: 02 jul. 2024.

FARIA, M. D. de; CARVALHO, A. F. dos S. de; CARVALHO, J. L. F. O Consumidor em ambientes de varejo online e offline: similitudes, contrastes e influências recíprocas. In: IX SIMPÓSIO EM EXCELÊNCIA DE GESTÃO E TECNOLOGIA - SEGET, Resende, 2012. **Anais...** IX Simpósio em Excelência de Gestão e Tecnologia – SEGET, 2012. Disponível em: <https://www.aedb.br/seget/arquivos/artigos12/21916169.pdf>. Acesso em: 31 mai. 2024.

FAVORETE, A. P. Z.; PEREIRA, A. T. Impactos da pandemia sobre o e-commerce. **Caderno PAIC**, v. 22, n. 1, p. 117-130, 2021.

FERRARA, G. H. **Uma análise da contribuição das redes sociais na Internet para o comércio eletrônico: social-commerce**. 2013. 180f. Dissertação (Mestrado em Tecnologia da Inteligência) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2013.

GALINARI, R.; CERVIERI JÚNIOR, O.; TEIXEIRA JÚNIOR, J. R.; RAWET, E. L. Comércio eletrônico, tecnologias móveis e mídias sociais no Brasil. **BNDES Setorial**, n. 41, p. 177-180, mar. 2015.

GÓMEZ-GALÁN, J.; MARTÍNEZ-LÓPEZ, J. A.; LÁZARO-PÉREZ, C.; SÁNCHEZ-SERRANO, J. L. S. Social Networks Consumption and Addiction in College Students during the COVID-19 Pandemic: Educational Approach to Responsible Use. 2020. **Sustainability**, v. 12, n.7737, 2020.

HAJLI, M. N. A study of the impact of social media on consumers. **International Journal of Market Research**, v. 56, n. 3, p. 387-404, 2014.

HINSCH, C.; SHELDON, K. The impact of frequent social Internet consumption: Increased procrastination and lower life satisfaction. **Journal of Consumer Behaviour**, v. 12, n. 6, p. 496-505, nov./dez. 2013.

HIRANO, K.; IMBENS, G. W. Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. **Health Services & Outcomes Research Methodology**, v. 2, p. 1-20, dez. 2001.

HUANG, W-L.; HU, P.; TSAI, S.; CHEN, X-D. The business analysis on the home-bias of E-commerce consumer behavior. **Electronic Commerce Research**, v. 21, p. 855-879, 2021.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. De 2010 a 2022, população brasileira cresce 6,5% e chega a 203,1 milhões. **IBGE**, Rio de Janeiro, 28 jun. 2023. Disponível em: <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/37237-de-2010-a-2022-populacao-brasileira-cresce-6-5-e-chega-a-203-1-milhoes>. Acesso em: 01 set. 2023.

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD) 2021**. Rio de Janeiro: IBGE, 2021.

IPEA - INSTITUTO DE PESQUISA ECONÔMICA APLICADA. **Vendas on-line no Brasil: uma análise do perfil dos usuários e da oferta pelo setor de comércio**. Comunicado Ipea nº 95. Brasília: Ipea, 2011.

LI, H.; KUO, C.; RUSSELL, M. The Impact of Perceived Channel Utilities, Shopping Orientations, and Demographics on the Consumer's Online Buying Behavior. **Journal of Computer-Mediated Communication**, v. 5, n. 2, dec. 1999.

LIMA-FILHO, D.; ALVES, C.; QUEVEDO-SILVA, F.; MOREIRA, L.; GARCEZ, V.; ARATANI, W. Profile of the electronic commerce consumer: a study with brazilian university students. **The Journal of Internet Banking and Commerce**, v. 17, p. 1-16, apr. 2012.

LINDA, S. L. A. I. Social commerce—e-commerce in social media context. **World Academy of Science Engineering and Technology**, v. 72, p. 39-44, dec. 2010.

LOHSE, G. L.; BELLMAN, S.; JOHNSON, E. J.; Consumer Buying Behavior on the Internet: Findings from Panel Data. **Journal of Interactive Marketing**, v. 14, n. 1, p. 15-29, 2000.

LONG, J. S.; FREESE, J. **Regression Models for Categorical Dependent Variables Using Stata**. 3th ed. Texas: Stata Press, 2014.

LOPEZ, S. F. M. **O comércio eletrônico como instrumento no processo de marketing de relacionamento: um estudo de multicaso**. 2003. 296f. Tese (Doutorado em Administração) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2003.

MACÊDO, K. A. B. **Ensaio sobre os efeitos da distribuição desigual do trabalho entre as famílias no Brasil**. 2022. 76 f. Dissertação (Mestrado em Economia) – Centro Acadêmico do Agreste, Universidade Federal de Pernambuco, Caruaru, 2022.

MADNI, G. Consumer's Behavior and Effectiveness of Social Media. **Global Journal of Management and Business Research**, v. 14, n. 8, 2014.

MAGALHÃES, R. A.; FERREIRA, D. R.; SILVA, T. L. A regulamentação do comércio eletrônico. **Revista Foco**, v. 15, n. 6, p. 1-18, 2022.

MARKUS, N. de P.; OLIVEIRA, F. R.; NISHIMURA, F. N. Efeito da educação pré-primária no desempenho escolar. In: XVIII ENCONTRO NACIONAL DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS REGIONAIS E URBANOS – XVIII ENABER, 2020, *on-line*. **Anais...** XVIII Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos – XVIII ENABER. Disponível em: https://brsa.org.br/wp-content/uploads/wpcf7-submissions/1863/Artigo_ENABER_identificado.pdf. Acesso em: 19 de fev. 2024.

MASTRANGELO, J. P.; MAIA, A. G. Os impactos da segurança da posse no desmatamento: novas evidências para a floresta amazônica. In: 49º Encontro Nacional de Economia - ANPEC, 2021, *on-line*. **Anais...** 49º Encontro Nacional de Economia - ANPEC, 2021. Disponível em: https://www.anpec.org.br/encontro/2021/submissao/files_l/i11-03f90977312377f030031306e64f4958.pdf. Acesso em: 31 mar. 2024.

McFADDEN, D. Conditional Logit Analysis of Qualitative Choice Behavior. In: ZAREMBKA, P. **Frontiers in econometrics**. New York: Academic Press 1974. Chap. 4th, pp. 105-142.

MORGADO, M. G. **Comportamento do consumidor online: perfil, uso da Internet e atitudes**. 2003. 160 f. Tese (Doutorado em Mercadologia) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, São Paulo, 2003.

NASCIMENTO, R. M. **E-commerce no Brasil: perfil do mercado e do e-consumidor brasileiro**. 2011. 77f. Dissertação (Master in International Management) – Escola Brasileira de Administração Pública e de Empresas, Fundação Getúlio Vargas, [s.l.], 2011.

NIC.br – NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR. **Pesquisa sobre o uso das tecnologias de informação e comunicação:** Pesquisa TIC Domicílios, ano 2015. [s.l.]: NIC.br, 2016. Disponível em: <http://cetic.br/pt/arquivos/domicilios/2015/individuos/>. Acesso em: 20 ago. 2023.

NIC.br – NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR. **Pesquisa sobre o uso das tecnologias de informação e comunicação:** Pesquisa TIC Domicílios, ano 2016. [s.l.]: NIC.br, 2017. Disponível em: <http://cetic.br/pt/arquivos/domicilios/2016/individuos/>. Acesso em: 20 ago. 2023.

NIC.br – NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR. **Pesquisa sobre o uso das tecnologias de informação e comunicação:** Pesquisa TIC Domicílios, ano 2018. [s.l.]: NIC.br, 2019. Disponível em: <http://cetic.br/pt/arquivos/domicilios/2018/individuos/>. Acesso em: 20 ago. 2023.

NIC.br – NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR. **Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros:** Pesquisa TIC Domicílios, ano 2019. [s.l.]: NIC.br, 2020. Disponível em: <http://cetic.br/pt/arquivos/domicilios/2019/individuos/>. Acesso em: 20 ago. 2023.

NIC.br – NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR.

Pesquisa sobre o uso das tecnologias de informação e comunicação nos

domicílios brasileiros: pesquisa TIC Domicílios (Edição COVID-19 - Metodologia adaptada), ano 2020. [s.l.]: NIC.br, 2021. Disponível em:

<<https://cetic.br/pt/arquivos/domicilios/2020/individuos/>>. Acesso em: 20 ago. 2023.

NIC.br – NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR.

Pesquisa sobre o uso das tecnologias de informação e comunicação nos

domicílios brasileiros: pesquisa TIC Domicílios, ano 2021. [s.l.]: NIC.br, 2022.

Disponível em: <https://cetic.br/pt/arquivos/domicilios/2021/individuos/>. Acesso em: 20 ago. 2023.

NIC.br – NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR.

Pesquisa sobre o uso das tecnologias de informação e comunicação nos

domicílios brasileiros: pesquisa TIC Domicílios, ano 2022. [s.l.]: NIC.br, 2023.

Disponível em: <https://cetic.br/pt/arquivos/domicilios/2022/individuos/>. Acesso em: 20 ago. 2023.

NIC.br – NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR..

Pesquisa sobre o uso das tecnologias de informação e comunicação: Pesquisa

TIC Domicílios, ano 2017. [s.l.]: NIC.br, 2018. Disponível em:

<http://cetic.br/pt/arquivos/domicilios/2017/individuos/>. Acesso em: 20 ago. 2023.

OLIVEIRA, C. A. de. **Dark kitchens:** uma análise econômica do modelo de negócios e de sua regulação. Brasília: Escola Nacional de Administração Pública, 2023.

OPINIONBOX. E-Commerce Trends 2024: Perspectivas sobre o futuro das vendas online a partir do comportamento do consumidor. [s.l.]: **Opinionbox**, 2023.

Disponível em: <https://www.commercetrends.com.br/>. Acesso em: 02 set. 2023.

PAUL, J.; BAKER, H.; COCHRAN, J. Effect of Online Social Networking on Student Academic Performance. **Computers in Human Behavior**, v. 28, n. 6, p. 2117-2127, nov. 2012.

PREMEBIDA, E. A. E-commerce em 2020, um cenário de oportunidades em meio a pandemia. **Research, Society and Development**, v. 10, n. 2, p. e59210212984-e59210212984, 2021.

ROBINS, J. M.; ROTNITZKY, A. Semiparametric efficiency in multivariate regression models with missing data. **Journal of the American Statistical Association**, v. 90, n. 429, p. 122-129, mar. 1995.

ROBINS, J. M.; ROTNITZKY, A; ZHAO, L. P. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. **Journal of the American Statistical Association**, v. 90, n. 429, p. 106-121, mar. 1995.

ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. **Biometrika**, v. 70, n. 1, p. 41-55, 1983.

RUBIN, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. **Journal of Educational Psychology**, v. 66 n. 5, p. 688-701, 1974.

RUBIN, D. B. Using propensity scores to help design observational studies: Application to the tobacco litigation. **Health Services and Outcomes Research Methodology**, v. 2, p. 169-188, dez. 2001.

RYBACZEWSKA, M.; SPARKS, L. Ageing consumers and e-commerce activities. **Ageing & Society**, v. 42, n. 8, p. 1879-1898, 2022.

SAHA, M.; SAHNEY, S. Exploring the relationships between socialization agents, social media communication, online shopping experience, and pre-purchase search: a moderated model. **Internet Research**, v. 32, p. 536-567, 2021.

SALGADO, D. Influencer digital: tudo que você precisa saber para fazer sucesso no *marketing* digital. **Opinion Box**, [s.l.], 29 ago. 2022. Disponível em: <https://blog.opinionbox.com/influencer-digital/#:~:text=O%20influencer%20digital%20%C3%A9%20uma,ela%20produz%20nas%20redes%20sociais>. Acesso em 15 nov. 2023.

SHAPIRO, C.; VARIAN, H. R. **El dominio de la información**: una guía estratégica para la economía de la red. Barcelona: Antoni Bosch, 1999.

SHI, J.; WANG, Y. Can Online Social Network Promote Residents' Consumption—Proven by the Micro Data from 27,632 Chinese Households. **Open Journal of Business and Management**, v. 8 n. 2, mar. 2020.

SILVA, W. M. da; MORAIS, L. A. de; FRADE, C. M.; PESSOA, M. F. *Marketing* digital, E-commerce e pandemia: uma revisão bibliográfica sobre o panorama brasileiro. **Research, Society and Development**, v. 10, n. 5, p. e45210515054-e45210515054, 2021.

SONG, S.; SUN, Q. Online Consumption and Income Efficiency: Evidence from China. **The Chinese Economy**, v. 53, n. 6, p. 465-476 jul. 2020.

SOURBATI, M. 'It could be useful, but not for me at the moment': older people, internet access and e-public service provision. **New Media & Society**, v. 11, n. 7, p. 1083-1100, 2009.

TAN, Z. Bounded, efficient and doubly robust estimation with inverse weighting. **Biometrika**, v. 97, n. 3, p. 661-682, 2010.

UZUNOĞLU, E.; KIP, S. M. Brand communication through digital influencers: Leveraging blogger engagement. **International journal of information management**, v. 34, n. 5, p. 592-602, 2014.

WANG, J.; OH, J. Factors Influencing Consumers' Continuous Purchase Intentions on TikTok: An Examination from the Uses and Gratifications (U&G) Theory Perspective. **Sustainability**, v. 15, n. 13, p. 10028, 2023.

WILCOX, K.; STEPHEN, A. Are Close Friends the Enemy? Online Social Networks, Self-Esteem, and Self-Control. **Journal of Consumer Research**, v. 40, n. 1, p. 90-103, nov. 2012.

WU, Y.; HUANG, H. Influence of Perceived Value on Consumers' Continuous Purchase Intention in Live-Streaming E-Commerce - Mediated by Consumer Trust. **Sustainability**, v. 15, n. 5, p. 4432, 2023.

YADAV, M.; RAHMAN, Z. Measuring consumer perception of social media *marketing* activities in e-commerce industry: **Scale development & validation. Telematics and Informatics**, v. 34, n. 7, p. 1294-1307, 2017.

ZHANG, Q.; YANG, J. Model construction of hierarchical polarization characteristics combined with social E-commerce consumer behavior. **Security and Communication Networks**, v. 2022, 2022.

ZHANG, S.; HUANG, C.; LI, X., REN, A. Characteristics and roles of streamers in e-commerce live streaming. **The Service Industries Journal**, v. 42, n. 13-14, p. 1001-1029, 2022.

ZHANG, Y.; TRUSOV, M.; STEPHEN, A.; JAMAL, Z. Online Shopping and Social Media: Friends or Foes? **Journal of Marketing**, n. 81, p. 24-41, 2017.

APÊNDICE

Apêndice A – Probit

Tabela 6 - Probit: Categoria de referência

Variável	Categoria de referência
Ano	2015
Sexo	Feminino
PEA	Na força de trabalho
Escolaridade	Analfabeto/Educação infantil
Renda familiar	Até 1 SM
Etnia	Branca
Área	Urbana
Região	Sudeste
Redes sociais	Não usou redes sociais
Transações financeiras	Não realizou transação financeira

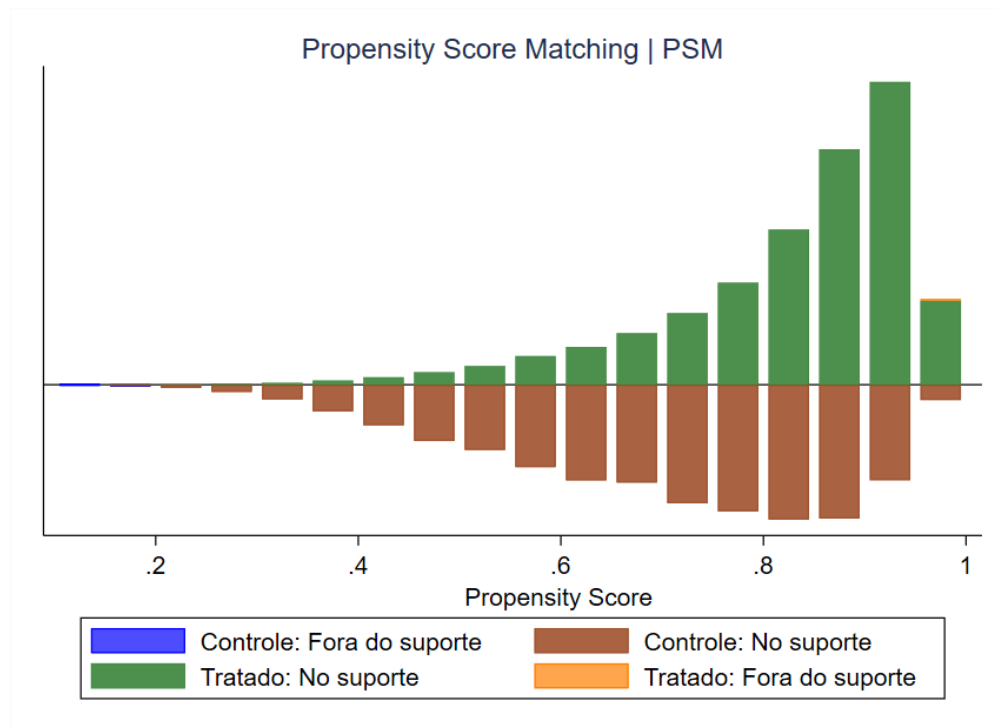
Fonte: própria autora (2024).

Apêndice B – *Propensity Score Matching* (PSM)

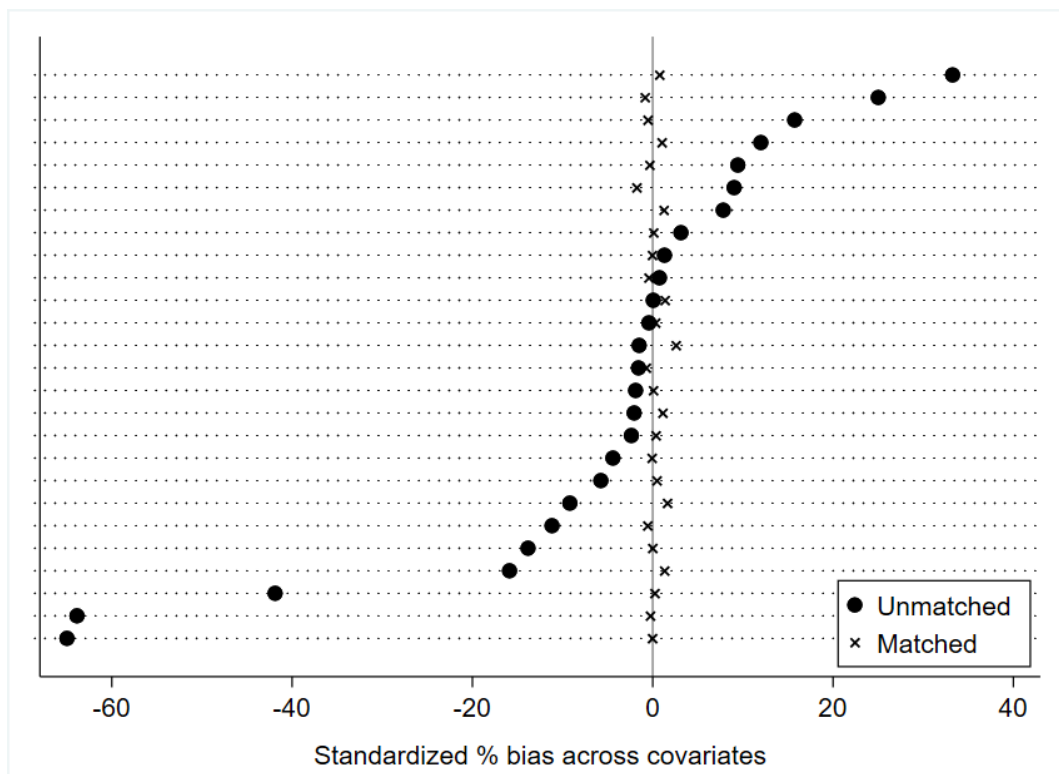
Tabela 7 - Modelo *Propensity Score Matching* - Logit

Variável dependente: Fez compras <i>on-line</i> no último ano			
	Coeficiente	Erro Padrão	P> z
Ano			
2016	0,0519	0,0432	0,229
2017	0,1048	0,0417	0,012
2018	0,1015	0,0409	0,013
2019	0,1793	0,0409	0,000
2021	0,5546	0,0421	0,000
2022	0,7206	0,0431	0,000
Sexo	-0,3930	0,0215	0,000
PEA	-0,2968	0,0270	0,000
Renda familiar			
Mais de 1 SM até 2 SM	0,2697	0,0258	0,000
Mais de 2 SM até 3 SM	0,4211	0,0318	0,000
Mais de 3 SM até 5 SM	0,5904	0,0382	0,000
Mais de 5 SM	0,5395	0,0458	0,000
Escolaridade			
Fundamental	0,5853	0,0499	0,000
Médio	1,1527	0,0497	0,000
Superior	1,6243	0,0565	0,000
Idade	-0,0856	0,0063	0,000
Idade²	0,0004	0,0004	0,000
Etnia			
Preta	-0,0179	0,0329	0,586
Parda	-0,0289	0,0250	0,248
Amarela	-0,0877	0,0643	0,173
Indígena	-0,0607	0,0795	0,445
Área	-0,3552	0,0329	0,000
Região			
Nordeste	-0,2026	0,0279	0,000
Sul	0,1697	0,0354	0,000
Norte	-0,3009	0,0334	0,000
Centro-Oeste	-0,0853	0,0376	0,023
Constante	3,4869	0,1470	0,000
Pseudo R2	0,1197		
Log Likelihood	-30197,279		
Observações	69.295		

Fonte: adaptação da tabela extraída do *Software Stata*®.

Gráfico 9 - Suporte Comum

Fonte: gráfico gerado através do *Software Stata®*.

Gráfico 10 - PSM Matching

Fonte: gráfico gerado através do *Software Stata®*.