

Universidade Federal do Paraná
Setor de Ciências Exatas
Departamento de Estatística
Programa de Especialização em *Data Science* e *Big Data*

Ihorana Aguilar Cuco

**Anulação Eleitoral: um olhar sobre as condições
socioeconômicas nas eleições a governador nos
municípios do Paraná em 2018**

**Curitiba
2022**

Ihorana Aguilar Cuco

**Anulação Eleitoral: um olhar sobre as condições
socioeconômicas nas eleições a governador nos
municípios do Paraná em 2018**

Monografia apresentada ao Programa de
Especialização em Data Science e Big Data da
Universidade Federal do Paraná como requisito
parcial para a obtenção do grau de especialista.

Orientador: Wagner Bonat

Curitiba
2022

Anulação Eleitoral: um olhar sobre as condições socioeconômicas nas eleições a governador nos municípios do Paraná em 2018

Ihorana Aguilar Cuco¹
Wagner Bonat²

Resumo

O debate sobre o comportamento alienado do eleitor tem sido um dos grandes desafios dentro da dimensão eleitoral além de ser um fenômeno crescente e relevante tanto nas democracias consolidadas como nas novas. Sendo assim, o presente trabalho buscou explorar variáveis que poderiam ter alguma associação com a anulação eleitoral (votos brancos e nulos) nos municípios do Paraná. A metodologia do estudo foi baseada em um Modelo de Regressão Linear Múltipla utilizando dados eleitorais de 2018 e dados secundários sobre indicadores socioeconômicos. O estudo procurou avaliar até que ponto as condições socioeconômicas estimulam, retraem ou interferem no comportamento de anular os votos dos eleitores. Os resultados reforçam o entendimento de que, os votos brancos para governador em 2018 podem estar associados a contextos de menor desenvolvimento educacional, ou seja, mais propícios ao erro não intencional dos eleitores. No caso dos votos nulos, isso não parece acontecer. Além disso, quanto mais despesas em investimentos e educação o município faz, há menor votos brancos e nulos.

Palavras-chave: Alienação eleitoral, Eleições.

Abstract

The debate on voter alienated behavior has been one of the great challenges within the electoral dimension, in addition to being a growing and relevant phenomenon both in consolidated and new democracies. Therefore, the present work sought to explore variables that could have some association with electoral annulment (white and null votes) in the municipalities of Paraná. The study methodology was based on a Multiple Linear Regression Model using 2018 electoral data and secondary data from socioeconomic indicators. The results reinforce the understanding that the white vote for governor in 2018 may be associated with contexts of lower educational development, that is, more prone to unintentional electoral error. In the case of null votes, this does not seem to happen. In addition, the

¹Aluno do programa de Especialização em Data Science & Big Data, ihoranacuco@gmail.com.

²Professor adjunto da Universidade Federal do Paraná e coordenador do Programa do curso de especialização em Data Science Big Data. wbonat@gmail.com.

more spending on investments and education the municipality makes, the fewer blank and null votes will be registered.

Keywords: Electoral Alienation, Elections.

1 Introdução

A não participação eleitoral (abstenção), somada àqueles eleitores que comparecem às urnas, mas optam por anular ou votar em branco, caracteriza o que tem sido denominado de alienação eleitoral. O presente trabalho tem como objetivo analisar somente os casos de votos nulos e brancos e se as bases sociais podem estar associadas a essas atitudes.

No Brasil, existem alguns trabalhos que abordam essa temática, mas ainda há grande esforço na literatura em entender os dois casos [1, 2, 3, 4]. O trabalho de (author?) [4], por exemplo, analisou separadamente abstenções, votos brancos e nulos nas eleições presidenciais de 1989 a 2002, utilizando dados agregados ao nível dos estados e concluiu que os efeitos do Índice de Desenvolvimento Humano (IDH) foram negativos tanto sobre brancos quanto sobre nulos.

Segundo os primeiros estudos sobre o tema, predominou o entendimento de que os votos inválidos poderiam representar um descontentamento de setores mais modernos da sociedade com o sistema político (author?) [5] ou um reflexo da falta de credibilidade do processo político (author?) [1]. Outras análises enfatizam que esses tipos de votos aumentam em contextos de menor alfabetização e extensão territorial (author?) [3] e em contextos de menor renda e alfabetização (author?) [2]. Dados do OCDE mostram que nas últimas eleições para as quais existem dados disponíveis, a participação eleitoral no Brasil foi de 80% dos eleitores cadastrados. Esse número é superior à média da OCDE, de 69%, e reflete a prática do voto obrigatório no Brasil. Porém, o status sócio econômico também pode afetar a participação eleitoral, dado que para os 20% mais favorecidos da população, esta taxa estimada é de 81% e, para os 20% menos favorecidos, de 79%. Esse dado mostra a relevância de entender os fatores sociais que de fato podem levar a alienação eleitoral (author?) [6].

2 Materiais e Métodos

Nesta seção será apresentado o conjunto de dados considerados no estudo e o método de Regressão Linear Múltipla que foi empregado nesse trabalho. Inicialmente, as covariáveis foram separadas por grupos socioeconômicos e para a estimação foi utilizado o método de mínimos quadrados, por ser uma técnica de otimização matemática que busca encontrar o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados (resíduos). O objetivo é encontrar os valores dos parâmetros para o modelo que "melhor" se ajusta aos dados.

2.1 O conjunto de dados

O conjunto de dados socioeconômicos utilizado é disponibilizado de forma pública pelo IPARDES. As variáveis dependentes do estudo são: Taxa de votos branco e taxa de votos nulos para a eleições de 2018 a Governador nos municípios do Paraná. Como para alguns municípios do Paraná não haviam informações de finanças municipais, foram considerados apenas municípios que tinham esses dados disponíveis. No final da análise, sobraram 285 municípios paranaenses. A seguir serão apresentadas as 24 variáveis que foram consideradas inicialmente no estudo com sua definição de acordo com o site do Instituto Paranaense de Desenvolvimento Econômico e Social (IPARDES).

- ▶ Despesas per capita: Total da despesa efetivamente realizada durante o exercício fiscal sobre o total da população estimada para 2018 (R\$).
- ▶ Despesas de transferências per capita: Recursos destinados pelo poder público a outras pessoas de direito público ou privado para o financiamento de suas despesas correntes sobre o total da população estimada para 2018 (R\$).
- ▶ Fundo de Participação Municipal: Fundo de participação dos municípios transferido pela União sobre o total da população estimada para 2018 (R\$)
- ▶ Despesas com educação e segurança: Participação percentual das despesas decorrentes das ações voltadas para educação, saúde, previdência e assistência social no total da despesa efetivamente realizada durante o exercício fiscal (R\$).
- ▶ Receita com educação per capita: Despesas decorrentes das ações voltadas para a educação sobre o total da população estimada para 2018 (R\$).
- ▶ Despesas com investimentos per capita: Despesas com investimentos sobre o total da população estimada para 2018 (R\$).
- ▶ Receita per capita: Receitas orçamentárias recolhidas aos cofres públicos por força de arrecadação sobre o total da população estimada para 2018 (R\$).
- ▶ Despesa com saúde per capita: Refere-se à despesa municipal total realizada com ações e serviços públicos de saúde sobre o total da população estimada para 2018 (R\$).
- ▶ Receita tributária per capita: Receita que corresponde à arrecadação da receita tributária menos as deduções sobre o total da população estimada para 2018 (R\$).
- ▶ IDHM educação: O Índice de Desenvolvimento Humano Municipal é medido pela composição de dois subíndices: a escolaridade da população adulta (% de pessoas de 18 anos ou mais de idade com ensino fundamental completo) e; o fluxo escolar da população jovem (% de crianças de 5 a 6 anos frequentando escola; % de jovens de 11 a 13 anos frequentando os anos finais do ensino fundamental; % de jovens de 15 a 17 anos com ensino fundamental completo e; % de jovens de 18 a 20 anos com ensino médio completo).
- ▶ IDHM renda: No Índice de Desenvolvimento Humano Municipal Renda, o domínio sobre recursos para garantir um padrão de vida com acesso a necessidades básicas, como água, alimento e moradia, é medido pela renda mensal per capita.
- ▶ IDHM alfabetização: É o percentual das pessoas acima de 15 anos de idade que são alfabetizadas, ou seja, que sabem ler e escrever pelo menos um bilhete simples.
- ▶ IDHM: O IDHM é um índice composto que agrega três das mais importantes dimensões do desenvolvimento humano: a oportunidade de viver uma vida longa e saudável (IDHM Saúde); de ter acesso a conhecimento (IDHM Educação) e ter um padrão de vida que garanta as necessidades básicas (IDHM Renda).
- ▶ IPDM: O Índice Iparades de Desempenho Municipal (IPDM) procura avaliar a situação dos municípios paranaenses, considerando, com igual ponderação, as três principais áreas de desenvolvimento econômico e social, a saber: a) emprego, renda e produção agropecuária; b) educação; e c) saúde.
- ▶ IPDM educação: Avalia a situação dos municípios paranaenses, com relação a: taxa de matrícula na educação infantil; taxa de abandono escolar (1ª a 4ª série / 1º a 5º ano; 5ª a 8ª série / 6º a 9º ano e ensino médio); taxa de distorção idade-série (1ª a 4ª série / 1º a 5º ano; 5ª a 8ª série / 6º a 9º ano e ensino médio); percentual de docentes com ensino superior (1ª a 4ª série / 1º a 5º ano; 5ª a 8ª série / 6º a 9º ano e ensino médio); resultado do IDEB (1ª a 4ª série / 1º a 5º ano e 5ª a 8ª série / 6º a 9º ano).
- ▶ IPDM renda: Na construção do índice da dimensão Emprego, Renda e Produção Agropecuária são utilizadas as variáveis relacionadas ao salário médio, ao emprego formal e à renda da agropecuária.
- ▶ IPDM saúde: Na construção do índice da dimensão Saúde são usadas as variáveis: número de consultas pré-natais; óbitos infantis por causas evitáveis, e óbitos por causas mal-definidas.
- ▶ PIB per capita: PIB dividido pela população estimada do município em 2018.

- ▶ Extremamente pobre: Proporção dos indivíduos com renda domiciliar per capita igual ou inferior a R\$ 70,00 (setenta reais) mensais.
- ▶ Vulneráveis a pobreza: Proporção de indivíduos com renda domiciliar per capita igual ou inferior a R\$ 255,00 (duzentos e cinquenta e cinco reais) mensais.
- ▶ Taxa de pobreza: Proporção dos indivíduos com renda domiciliar per capita igual ou inferior a R\$ 140,00 mensais, em reais de agosto de 2010. O universo de indivíduos é limitado àqueles que vivem em domicílios particulares permanentes.
- ▶ Eleitores analfabetos: Proporção de pessoas analfabetas sobre o total de pessoa a quem se atribui a aptidão de participar da escolha de um representante para exercer cargo público.
- ▶ Eleitores femininos: Proporção de pessoas do sexo feminino sobre o total de pessoa a quem se atribui a aptidão de participar da escolha de um representante para exercer cargo público.
- ▶ Eleitores masculinos: Proporção de pessoas do sexo masculino sobre o total de pessoa a quem se atribui a aptidão de participar da escolha de um representante para exercer cargo público.

A partir dos dados listados, foram feitas análises descritivas de cada variável resposta com todas as variáveis explicativas. Como são muitas variáveis, nesse estudo foram destacados apenas os gráficos que mostraram alguma associação com as variáveis resposta taxa de votos nulos e taxa de votos brancos. É possível observar que, no caso dos votos em branco, parece haver uma associação negativa com as variáveis de finanças públicas e, positiva com as variáveis sociais.

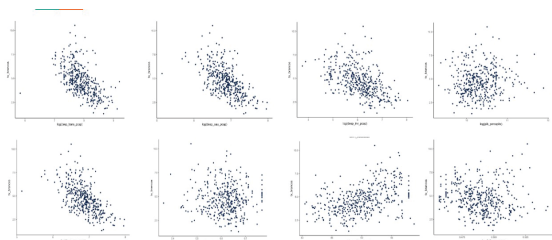


Figura 1: Gráfico de dispersão Modelo 1 - Variável resposta taxa de voto branco

Para o Modelo 2, também parece haver uma associação negativa dos votos nulos com as variáveis de finanças municipais e positiva com as variáveis sociais de educação e alfabetização.

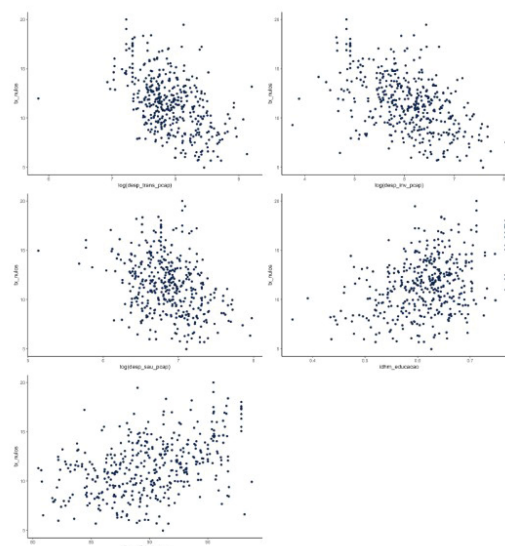


Figura 2: Gráfico de dispersão Modelo 2 - Variável resposta taxa de voto nulo

2.2 Modelos empregados

O Modelo empregado foi o método de Regressão Linear Múltipla. Serão considerados dois modelos, um para explicar as taxas de votos nulos e o outro para explicar as taxa de votos em branco. O modelo será apresentado no tópico 2.2.1.

2.2.1 Regressão Linear Múltipla

A Regressão Linear Múltipla é um modelo de análise que se usa quando se deseja modelar a relação linear entre uma variável de desfecho contínua e múltiplas variáveis preditoras que podem ser contínuas ou categóricas. Onde Y é a variável de interesse, a que queremos prever. x é a variável preditora e que está associado ao beta, que dirá o quão inclinado está a reta, ou o quão influente é a variável preditora. Por sua vez, o alfa é o valor que descreve o intercepto: onde a linha está quando o valor de x é zero.

$$Y = \alpha + \beta_1 x_1 + \dots + \beta_2 x_2 + e$$

3 Resultados e Discussões

Para cada grupo das 24 covariáveis consideradas (finanças públicas, IDHM, IPDM, sociais e renda) do conjunto de dados, foi aplicado o método de stepwise para identificar um subconjunto útil de preditores. O critério utilizado foi o de Akaike (AIC), a fim de mensurar a qualidade do modelo estatístico. Nesse caso, menores valores de AIC representam uma maior qualidade e simplicidade. Finalmente, para cada modelo, foram analisados os pressupostos.

3.1 Ajuste dos modelos

Para o ajuste do modelo, foi feito o método de stepwise. O processo adiciona sistematicamente a variável mais significativa ou remove a variável menos significativa durante cada etapa e o melhor modelo é escolhido com base no menor AIC. Abaixo então os melhores modelos escolhidos para a variável resposta taxa de votos em branco e taxa de votos nulos. O Modelo 1 é referente a variável resposta taxa de branco e o Modelo 2 é referente a variável resposta taxa de nulo.

```
Step: AIC=75.62
tx_branco ~ log(desp_trans_pcap) + log(desp_inv_pcap) + log(desp_sau_pcap) +
idhm_educacao + idhm_alfabetizacao + log(pib_percapita) +
tx_eleitores_masc

Df Sum of Sq  RSS  AIC
<none>                351.31  75.617
- log(pib_percapita)  1    3.788 355.10  76.674
+ ipdm_renda          1    1.137 350.17  76.694
+ ipdm_educacao       1    0.360 350.95  77.325
+ tx_eleitores_analf  1    0.304 351.01  77.371
+ log(rec_trib_pcap)  1    0.148 351.16  77.497
+ ipdm_saude          1    0.080 351.23  77.552
- log(desp_sau_pcap)  1    6.638 357.95  78.953
- tx_eleitores_masc   1    6.703 358.01  79.004
- log(desp_trans_pcap) 1    7.068 358.38  79.295
- idhm_educacao       1   11.281 362.59  82.625
- log(desp_inv_pcap)  1   11.856 363.17  83.077
- idhm_alfabetizacao  1   45.470 396.78 108.305

MODEL INFO:
Observations: 285
Dependent Variable: tx_branco
Type: OLS linear regression
```

Figura 3: Modelo 1 selecionado pelo critério de AIC - Variável resposta taxa de votos brancos

```
Step: AIC=469.26
tx_nulos ~ log(desp_trans_pcap) + log(desp_inv_pcap) + log(desp_sau_pcap) +
idhm_educacao + idhm_alfabetizacao

Df Sum of Sq  RSS  AIC
<none>                1417.9 469.26
- log(desp_sau_pcap)  1   10.231 1428.1 469.31
+ ipdm_renda          1    5.216 1412.7 470.21
+ tx_eleitores_masc   1    4.036 1413.8 470.45
+ ipdm_educacao       1    1.291 1416.6 471.00
+ log(rec_trib_pcap)  1    1.211 1416.7 471.02
+ tx_eleitores_analf  1    1.024 1416.9 471.06
+ ipdm_saude          1    0.536 1417.3 471.15
+ log(pib_percapita)  1    0.000 1417.9 471.26
- idhm_alfabetizacao  1   23.849 1441.7 472.02
- idhm_educacao       1   52.233 1470.1 477.57
- log(desp_trans_pcap) 1   52.459 1470.3 477.62
- log(desp_inv_pcap)  1   96.887 1514.8 486.10

MODEL INFO:
Observations: 285
Dependent Variable: tx_nulos
Type: OLS linear regression
```

Figura 4: Modelo 2 selecionado pelo critério de AIC - Variável resposta taxa de votos nulos

A figura abaixo mostra as análises dos pressupostos dos dois modelos. Os primeiros quatro gráficos são referentes ao Modelo 1 e os demais ao Modelo 2. Os gráficos dos resíduos pelos valores previstos permite analisar tanto a linearidade quanto a homoscedasticidade. Como a linha vermelha está aproximadamente horizontal, considerou-se que o pressuposto de linearidade foi atendido. Os gráficos de normalidade permitem ver se os resíduos têm distribuição normal. Como os resíduos estão bastante em cima da linha pontilhada, considerou-se que está atendendo o pressuposto. Já os gráficos de homoscedasticidade, onde a linha vermelha

deve ser aproximadamente horizontal, também atendeu ao pressuposto e também percebe-se que não existem pontos de alavancagem.

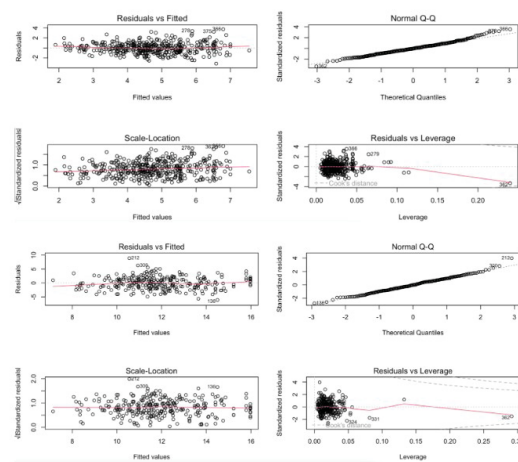


Figura 5: Gráficos de pressupostos dos modelos

3.2 Resultados

Abaixo estão os resultados da regressão linear múltipla dos Modelos 1 e 2.

tx_branco					tx_nulos				
Predictors	Estimates	CI	p		Predictors	Estimates	CI	p	
(Intercept)	-3.04	-11.31 - 5.23	0.470		(Intercept)	13.43	2.16 - 24.69	0.020	
desp_trans pcap [log]	-0.76	-1.40 - -0.13	0.019		desp_trans pcap [log]	-1.92	-3.09 - -0.74	0.001	
desp_inv pcap [log]	-0.41	-0.68 - -0.15	0.002		desp_inv pcap [log]	-1.18	-1.71 - -0.65	<0.001	
desp_sau pcap [log]	-0.73	-1.36 - -0.10	0.023		desp_sau pcap [log]	0.90	-0.35 - 2.16	0.157	
idhm_educacao	-5.11	-8.49 - -1.74	0.003		idhm_educacao	8.63	3.33 - 13.92	0.002	
idhm_alfabetizacao	0.15	0.10 - 0.19	<0.001		idhm_alfabetizacao	0.10	0.01 - 0.19	0.031	
pib_percapita [log]	0.36	-0.05 - 0.78	0.085		Observations	285			
tx_eleitores_masc	15.71	2.26 - 29.15	0.022		R ² / R ² adjusted	0.392 / 0.381			
Observations	285								
R ² / R ² adjusted	0.448 / 0.434								

Figura 6: Resultado do Modelo de Regressão Linear Múltipla

Conforme o R-quadrado de cada modelo, o conjunto de variáveis independentes escolhido explica melhor as variações das taxas de votos brancos nas eleições para Governador de 2018 para os municípios do Paraná em comparação com votos nulos.

No primeiro modelo, a variável do índice de desenvolvimento humano para alfabetização é a mais significativa e mostra que existe uma associação positiva entre as taxas de votos brancos e a alfabetização. No 2 modelo, a variável mais significativa é despesa de investimento, que mostra uma associação negativa com as taxas de votos nulos.

De forma geral, todas as variáveis de finanças públicas apresentam uma associação negativa com votos brancos e nulos.

A maioria das variáveis de renda e sociais apresenta uma associação positiva com votos brancos e nulos, apenas educação para o modelo 1 apresentou associação negativa.

Os resultados reforçam o entendimento de que, os votos

brancos para governador em 2018 podem estar associados a contextos de menor desenvolvimento educacional, ou seja, mais propícios ao erro não intencional dos eleitores. No caso dos votos nulos, isso não parece acontecer. Além disso, quanto mais despesas em investimentos e educação o município faz, há menor votos brancos e nulos.

4 Conclusões

Em ambos os modelos, percebe-se que quando os gastos com transferências municipais ou investimentos aumentam, espera-se uma variação negativa na taxa de votos nulos e em branco e que pode-se inferir que há uma atitude mais positiva dos cidadãos a projetos públicos que exijam esforços importantes de investimento, desde que entendam que suas condições de vida melhorarão. Esse fato confirma o resultado do modelo 2 que mostra que quando os gastos com saúde municipais aumentam 1% espera-se uma variação negativa na taxa de votos brancos.

Também observou-se que, quando os gastos com alfabetização aumentam espera-se uma variação positiva na taxa de votos brancos e nulos. Nessa mesma linha, quando olhamos para educação, que é um indicador mais rigoroso, ou seja, ele considera: a porcentagem de pessoas de 18 anos ou mais de idade com ensino fundamental completo, a porcentagem de crianças de 5 a 6 anos frequentando escola, a porcentagem de jovens de 11 a 13 anos frequentando os anos finais do ensino fundamental, a porcentagem de jovens de 15 a 17 anos com ensino fundamental completo e a porcentagem de jovens de 18 a 20 anos com ensino médio completo. Percebe-se que os efeitos da educação sobre votar branco e nulo mudam. Nesse caso, o efeito do voto nulo é negativo e do voto em branco positivo.

A taxa de eleitores analfabetos confirma que, municípios que possuem uma população melhor instruída tendem a ter mais votos em branco e a último variável, mostra que quanto mais eleitores do sexo masculino votam, aumenta a chance de votos em brancos.

E, por último, uma forma de aprimorar o estudo seria considerar análise espacial para a observabilidade de padrão nas votações de brancos e nulos entre os municípios.

Além disso, incluir um comparativo com eleições passadas para entender se a significância, a intensidade ou a relação das covariáveis mudou

Agradecimentos

Ao professor Wagner Bonat, que durante esse período contribui tanto para meu crescimento científico como intelectual. À equipe de professores do DSBD que trouxe uma abordagem abrangente das ferramentas e habilidades que um cientista de dados precisa. Ao meu marido, Luiz, pela grande parceria nessa jornada.

Referências

- [1] Wanderley Guilherme dos Santos. *Crise e castigo: partidos e gerais na política brasileira*. São Paulo: Vértice 134, 1987.
- [2] Antonio Carlos Alkmin dos Reis. *A participação política e seus correlatos sócio-econômicos*. Sistema eleitoral brasileiro: teoria e prática. Rio de Janeiro: Rio Fundo Editora e IUPERJ, 1991.
- [3] Lima Júnior. *Democracia e instituições políticas no Brasil dos anos 80*. São Paulo: Loyola, 1993.
- [4] Homero de Oliveira Costa. *Democracia e representação política no Brasil: uma análise das eleições presidenciais (1989-2002)*. Porto Alegre/Natal: Sulina/Ed. UFRN, 2007.
- [5] Simon Schwartzman. *Estudo Comparado de sistemas burocráticos: conceitos e perspectivas*. Revista de Administração Pública 4.1 (1970): 79-a., 1970.
- [6] OCDE. *Engajamento cívico*. OCDE Better Life Index, 2022.