



LARC-PCS-EPUSP

Comentários sobre alegações infundadas ou falsas sobre as urnas eletrônicas brasileiras

Autores (em ordem alfabética):

- **Felipe Kenzo Shiraishi**, Engenheiro e Aluno de Mestrado na Escola Politécnica da Universidade de São Paulo (USP)
- **Lucas Lago**, Mestre em Engenharia da Computação pela Escola Politécnica da Universidade de São Paulo (USP)
- **Marcos Antonio Simplicio Junior**, Professor Associado da Escola Politécnica da Universidade de São Paulo (USP)
- **Paulo Matias**, Professor Adjunto do Departamento de Computação da Universidade Federal de São Carlos (UFSCar)
- **Tiago Barbin Batalhão**, Doutor em Física pela Universidade Federal do ABC (UFABC)
- **Wilson Vicente Ruggiero**, Professor Titular da Escola Politécnica da Universidade de São Paulo (USP)

11 de Novembro de 2022

Resumo Executivo

O presente documento tem por objetivo elucidar algumas dúvidas que surgiram recentemente sobre o processo eleitoral brasileiro, em particular devido a algumas das alegações feitas no documento intitulado “Relatório preliminar de análise das urnas eletrônicas usadas na eleição presidencial do Brasil no primeiro turno – 02 de outubro de 2022”, de autoria desconhecida, mas amplamente divulgado na Internet. As análises foram realizadas por pesquisadores da Universidade de São Paulo e UFSCar, todos com expertise na área de computação e cibersegurança e também experiências prévias com análise do ecossistema da urna eletrônica brasileira, concentrando-se em três **alegações falsas ou que carecem de qualquer embasamento ou fundamento**:

- A. A alegação de que “As urnas de modelo anterior a 2020 não têm qualquer documentação de auditoria recente, e relatórios anteriores referem não serem passíveis de auditoria. Não há documentação comprobatória acerca dos modelos 2009/2010/2011/2013/2015”, justificativa usada no documento para comparar “o modelo 2020 (auditado) versus os demais”.
- B. A observação de que há “diferenças sutis entre os entre os arquivos de ‘log’ (...) das urnas” o levando à alegação de que “há, confirmadamente, ao menos dois softwares nas urnas, em diferentes modelos, não dependentes de determinado modelo”, pois supostamente “Jamais poderia haver sequer uma diferença nesta sequência” e “Nada mais explica essa diferença que não ao menos duas versões de softwares”
- C. A existência de uma suposta “trava que não deixa a soma [do total de votos lançados para os candidatos Lula e Bolsonaro] ultrapassar determinado número”, o que leva a observações como “Modelos não-2020 têm ângulo fixo ‘máximo’, do qual os votos do Bolsonaro (ou do Lula) ‘não podem passar’. Nesta ‘faixa limite’, a soma dos votos Lula+Bolsonaro é fixa: 300+0, 200+100 ou 100+200, por exemplo. Este jamais seria um comportamento esperado.”

Cabe notar que os dois primeiros aspectos recaem sobre a área de expertise dos autores, enquanto o último tem muito mais relação com matemática básica e, portanto, não requer conhecimentos técnicos especializados.

Em suma, as conclusões da análise foram:

- **Alegação A: não tem qualquer embasamento**, dada a existência de análises de segurança bastante recentes tanto para as urnas modelo 2020 e 2015, enquanto modelos mais antigos foram alvo de análises similares no passado
- **Alegação B**: embora seja levantada uma possibilidade interessante a partir da observação realizada, essa **hipótese de existência de dois códigos fontes revela-se falsa**, já que foi possível reproduzir em laboratório o mesmo comportamento observado com um **único código fonte** (a saber, o próprio código fonte da urna eletrônica)
- **Alegação C: a afirmação de suposta estranheza nos dados carece de qualquer fundamento**, pois o gráfico obtido é um resultado natural do fato de que cada urna tem um número máximo de eleitores aptos a votar (em geral, da ordem de 400 eleitores, podendo chegar a cerca de 500 em alguns casos).

Este relatório tem teor técnico, mas tenta na medida do possível trazer para uma linguagem mais próxima do público em geral as observações realizadas, buscando demonstrar claramente as conclusões apresentadas.

Introdução

O TSE e a USP firmaram o Convênio 14/2021, com a finalidade de permitir ao Laboratório de Arquitetura e Redes de Computadores (LARC) do Departamento de Engenharia de Computação e Sistemas Digitais (PCS) da Escola Politécnica da Universidade de São Paulo (USP), planejar e executar testes de segurança sobre as urnas eletrônicas brasileiras. Como parte desse convênio, foram disponibilizadas para a universidade de São Paulo: duas unidades do modelo UE2015 e três unidades do modelo UE2020; a documentação correspondente ao ecossistema das urnas; e os códigos fontes e respectivos códigos compilados para realizar a carga das urnas para fins de testes.

Como essa parceria é de conhecimento público, membros do LARC-USP têm mantido um canal de comunicação aberto com diversas pessoas e entidades interessadas no tema “segurança das urnas eletrônicas”, incluindo parceiros em universidades no Brasil (e.g., a Universidade Federal de São Carlos - UFSCar) e no exterior. Em uma dessas interações, no dia 02/Novembro/2022, chegou ao conhecimento de pesquisadores da USP e da UFSCar o documento intitulado “Relatório preliminar de análise das urnas eletrônicas usadas na eleição presidencial do Brasil no primeiro turno – 02 de outubro de 2022”. Embora de autoria desconhecida, o documento fazia alegações diversas sobre supostas fraudes nas eleições de 2022, com base em dados públicos disponibilizados pelo TSE. Pouco tempo depois, foi realizada uma apresentação desse relatório por meio de uma live, por alguém identificado pela mídia como o argentino Fernando Cerimedo. Por essa razão, o relatório (ainda de autoria desconhecida) ficou conhecido em alguns meios como o “Relatório Argentino”. Por comodidade, esse relatório é também assim referenciado neste documento.

Embora o teor do Relatório Argentino parecesse duvidoso desde o princípio, decidiu-se investir algum tempo na sua análise, com o objetivo de confirmar ou refutar as alegações ali feitas de forma independente. Em particular, considerando a expertise dos pesquisadores na área de computação e cibersegurança, bem como experiências prévias com análise da urna eletrônica brasileira, as análises iniciais buscaram se concentrar em 2 aspectos principais:

- A. A alegação de que “As urnas de modelo anterior a 2020 não têm qualquer documentação de auditoria recente, e relatórios anteriores referem não serem passíveis de auditoria. Não há documentação comprobatória acerca dos modelos 2009/2010/2011/2013/2015”, justificativa usada no documento para comparar “o modelo 2020 (auditado) versus os demais”
- B. A observação de que há “diferenças sutis entre os entre os arquivos de ‘log’ (...) das urnas” o levando à alegação de que “há, confirmadamente, ao menos dois softwares nas urnas, em diferentes modelos, não dependentes de determinado modelo”, pois supostamente “Jamais poderia haver sequer uma diferença nesta sequência” e “Nada mais explica essa diferença que não ao menos duas versões de softwares”.

Assim, em princípio, decidiu-se não abordar as discussões que, ao menos em teoria, envolveriam análises estatísticas de dados. A razão para isso é dupla. A primeira é que, embora estatística seja parte da formação dos pesquisadores, esta não é exatamente a sua área de expertise. A segunda é que várias reportagens na mídia já analisaram o perfil demográfico de ao menos parte das 147 seções eleitorais em que um único

candidato recebeu votos, o que corresponde a 4 casos para Bolsonaro (totalizando 124 votos) e 143 casos para Lula (somando 16.579 votos); o que se observou nessas análise é que a unanimidade para o candidato Lula, considerada improvável por alguns grupos, em geral ocorreu em aldeias indígenas, comunidades ou povoados: conforme reportagem em <https://g1.globo.com/politica/eleicoes/2022/eleicao-em-numeros/noticia/2022/11/07/secoes-em-que-so-um-candidato-a-presidente-foi-votado-somam-001percent-dos-votos-validos-no-2o-turno.ghtml>, “em 22 delas , o local de votação tem ‘aldeia’ no nome; em 27, tem ‘comunidade’ e, em 51, tem ‘povoado’”. Argumentar se esse comportamento amplamente favorável ao candidato Lula era ou não o esperado de tais grupos não faz parte do escopo desta análise.

Após a análise dos pontos acima, decidiu-se então incluir um novo ponto à análise, embora, ao menos em teoria, ele fizesse parte da porção “estatística” do Relatório Argentino. Especificamente, como muitas das observações nessa porção envolvem equívocos de matemática básica na interpretação de gráficos (algo que, portanto, não exige qualquer conhecimento de estatística), decidiu-se por analisar uma alegação adicional:

- C. A existência de uma suposta “trava que não deixa a soma [do total de votos lançados para os candidatos Lula e Bolsonaro] ultrapassar determinado número”, o que leva a afirmações como “Modelos não-2020 têm ângulo fixo ‘máximo’, do qual os votos do Bolsonaro (ou do Lula) ‘não podem passar’. Nesta ‘faixa limite’, a soma dos votos Lula+Bolsonaro é fixa: 300+0, 200+100 ou 100+200, por exemplo. Este jamais seria um comportamento esperado.”

A seguir, são apresentados os resultados das análises de cada um desses pontos.

Alegação A: Sobre documentações de auditoria dos diferentes modelos de urnas

Como parte da parceria entre USP e TSE, membros do LARC-USP acompanharam os Testes Públicos de Segurança (TPS) realizados no período de 22 a 27 de novembro de 2021, ocasião em que um grupo de 26 investigadores (incluindo membros da Polícia Federal que costumemente participam do evento) puderam realizar testes de segurança diversos sobre urnas do modelo UE2015. Acompanhamento similar ocorreu no Teste de Confirmação, realizado de 11 a 13 maio de 2022, em que 5 grupos de investigadores que obtiveram algum sucesso durante o TPS voltaram ao TSE para repetir os ataques e avaliar as contramedidas propostas. Detalhes sobre o evento, incluindo os relatórios correspondentes, encontram-se publicamente disponíveis em <<https://www.justicaeleitoral.jus.br/tps/>>.

Em paralelo às análises realizadas por investigadores no TPS 2021, e mesmo antes da realização do evento, pesquisadores do LARC-USP já tinham tido a oportunidade de avaliar a segurança do modelo UE2015, inclusive propondo melhorias ao sistema por meio de relatórios internos ao convênio entre USP e TSE. Se apoiando nos conhecimentos adquiridos em tais testes, e tendo em vista que as urnas eletrônicas do modelo UE2020 não haviam sido disponibilizadas pelo fabricante para serem avaliadas no período estabelecido do TPS 2021, foi solicitado ao LARC-USP que fizesse uma análise de segurança e relatório público também para a UE2020. As atividades realizadas nesse sentido foram planejadas para serem compatíveis com os padrões dos TPS, e dentro da disponibilidade de tempo que atendesse as eleições do calendário 2022. O relatório dessa atividade foi entregue em Agosto de 2022, ocasião na qual o resumo executivo correspondente foi publicado juntamente com análises independentes realizadas por pesquisadores da Universidade de Campinas (Unicamp) e da Universidade Federal de Pernambuco (UFPE). Esses relatórios encontram-se publicamente disponíveis em <<https://www.tse.jus.br/comunicacao/noticias/2022/Agosto/universidades-validam-nova-urna-e-codigos-fonte-dos-sistemas-eleitorais-357621>>.

Ainda, toda eleição conta com os chamados “Testes de Integridade”, durante os quais as urnas sorteadas são colocadas em um ambiente de votação simulada. Durante esse teste, são lançados diversos votos nas urnas, sem qualquer proteção de sigilo, e então confere-se se o resultado obtido nos boletins de urna correspondem aos votos efetivamente lançados. Os votos em questão não vão para a totalização oficial, mas servem apenas para verificar que as urnas testadas não estão de alguma forma “desviando votos”. Contanto que a simulação seja próxima o suficiente de um ambiente de votação real, seria difícil para um “software fraudador” saber que está sob testes, e então um eventual desvio de votos poderia ser detectado. Como os Testes de Integridade acontecem em todos os Estados do Brasil, e com urnas selecionadas aleatoriamente, eles acabam cobrindo diversos dos modelos de urna sendo utilizados naquela eleição. De fato, uma equipe da USP acompanhou em 2022 os Testes de Integridade realizados na cidade de São Paulo durante os dois turnos da eleição, ambos ocorridos no Centro Cultural São Paulo e na Universidade Paulista (Unip) do Paraíso, e observou diferentes modelos sob testes. A Figura A1 ilustra esse fato para os testes ocorridos no Centro Cultural, no 2o turno das eleições de 2022. Nota: essa imagem foi

escolhida por mostrar claramente os dois modelos em bancadas vizinhas, facilitando a visualização; imagens de outros lugares do país também permitem constatar essa variedade, em ambos os turnos.

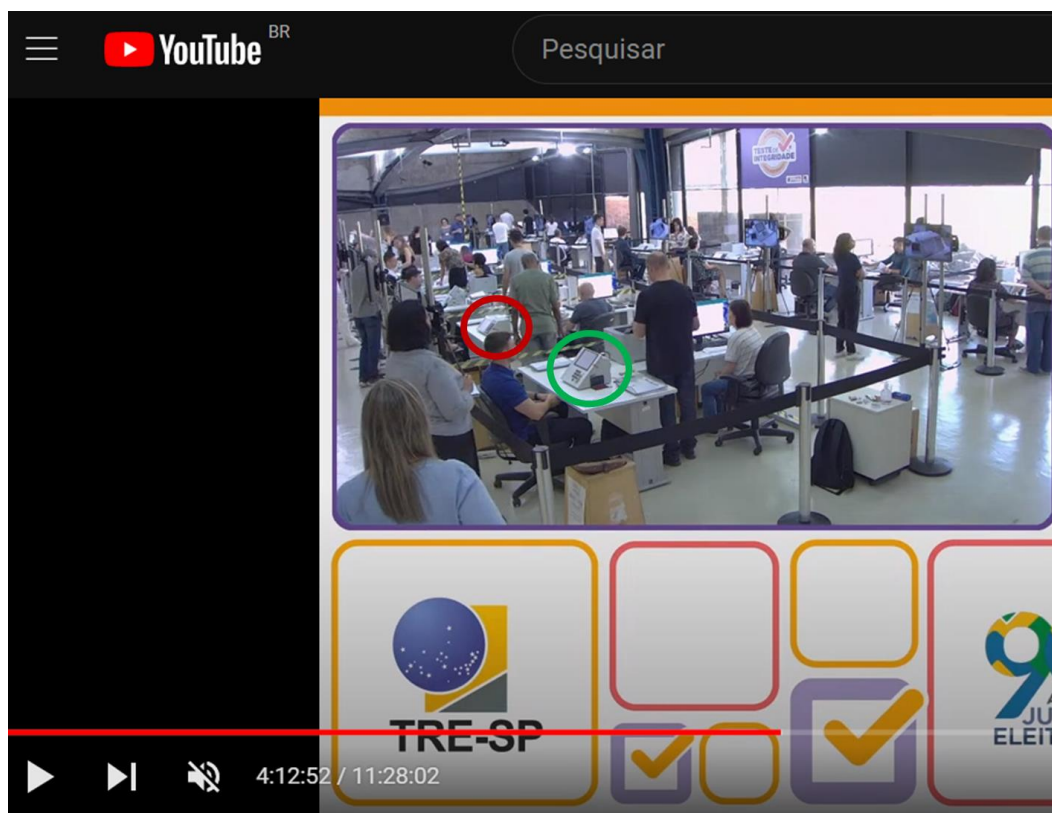


Figura A1 – Imagem do Teste Público de Integridade no 2o turno das Eleições de 2022, realizado no Centro Cultural de São Paulo. Pode-se observar o uso de urnas modelo 2020 (verde) e de um modelo anterior (vermelho), aproveitando a diferença visual entre elas. Fonte:

<https://www.youtube.com/watch?v=DpoTL7Nbhyyo>

Dado esse contexto, a análise da afirmação de que “As urnas de modelo anterior a 2020 não têm qualquer documentação de auditoria recente”, enquanto as urnas modelo 2020 teriam sido auditadas, parece exigir uma breve discussão sobre o que seria uma “auditoria”. O dicionário Michaelis define o termo de forma bem ampla, como o “Procedimento de análise, investigação e validação de um sistema, atividade ou informação”. Comumente, entretanto, o termo “auditoria” é reservado no meio técnico para análises de um sistema após seu uso, para avaliar alguma denúncia ou suspeita, por exemplo. Logo, a rigor, as análises de segurança feitas nos TPS e pelas equipes de universidades não poderiam ser consideradas “auditorias”, por terem acontecido antes das eleições apenas. Além disso, pode-se argumentar que os Testes de Integridade que ocorrem durante as eleições também não entrariam estritamente na definição de “auditoria”, por ocorrerem durante (e não após) as votações. Nesse sentido, até um pouco pedante, provavelmente seria classificado como “auditoria” nas eleições de 2022 somente análises como aquela feita pelo TCU <<https://portal.tcu.gov.br/imprensa/noticias/tcu-finaliza-analise-de-boletins-de-urna-do-1-turno-das-eleicoes.htm>> sobre 4.161 boletins de urna, em busca de divergências entre o resultado divulgado pelo TSE e o fornecido pelas urnas. Porém, nesse caso não se trataria de uma auditoria de urnas per se, mas sim do processo de totalização.

Isso posto, é difícil saber exatamente o que seria considerada uma “auditoria” pelo Relatório Argentino (nenhuma referência é ali fornecida para a mencionada “auditoria das urnas modelo 2020”). Portanto, assume-se aqui que o autor utiliza a definição mais ampla, de modo que tanto o TPS como o Teste de Integridade são considerados exemplos de auditoria. Se não for essa a definição utilizada, provavelmente seria necessário argumentar (novamente, de forma pedante), que nenhum modelo de urna passou por auditoria no contexto das eleições de 2022.

Dado esse esclarecimento/simplificação, não parece haver qualquer embasamento técnico na comparação do modelo de urna 2020 com modelos anteriores: a afirmação de que “As urnas de modelo anterior a 2020 não têm qualquer documentação de auditoria recente” é incorreta. Afinal, os relatórios públicos do TPS de 2021 testaram a segurança das urnas do modelo 2015, enquanto modelos anteriores à UE2015 foram analisados em edições anteriores dos Testes Públicos de Segurança, que normalmente ocorrem a cada 2 anos. Ao mesmo tempo, tanto a UE2015 como a UE2020 foram e continuam sendo analisadas pela USP, e relatórios deste último modelo foram publicados não apenas pela USP, mas também pela Unicamp e UFPE. Já os Testes de Integridade cobrem modelos diversos de urnas, via sorteio. Finalmente, ao afirmar que “relatórios anteriores referem não serem passíveis de auditoria”, o autor parece se referir ao Relatório de Auditoria das Eleições de 2014, disponível em <http://www.brunazo.eng.br/voto-e/arquivos/RelatorioAuditoriaEleicao2014-PSDB.pdf>. As conclusões desse relatório incluem essencialmente (1) uma listagem de pontos que puderam ser auditados, e para os quais não foram encontrados indícios de fraude, e (2) uma listagem de pontos que não puderam ser auditados a contento, com correspondentes sugestões de melhoria para sanar as lacunas observadas. Em nenhum momento o Relatório de Auditoria das Eleições de 2014 faz (ou sequer tenta fazer) qualquer distinção entre modelos de urnas, exceto para discutir características técnicas como a presença de hardware de segurança (presente em todas as urnas posteriores ao modelo 2009, ou seja, em todo o parque de urnas utilizadas nas eleições de 2022), e de mecanismo de biometria (o qual se encontrava e ainda se encontra disponível em vários modelos de urna). Portanto, novamente, não se justifica a divisão artificial entre modelos de urna feita no relatório sob análise.

Desta forma, do ponto de vista técnico, é difícil entender a razão pela qual o autor do Relatório Argentino decide fazer uma comparação entre [urnas modelo 2020] vs. [urnas mais antigas] como premissa das análises, assumindo (erroneamente) haver uma diferença clara da confiabilidade de cada modelo. Consequentemente, é bem provável que quaisquer conclusões tiradas a partir dessa comparação não tenham relação direta com o modelo em si, contrariamente àquilo que o relatório sob análise dá a entender. Por exemplo, ao menos em alguns casos, pode haver correlação entre o modelo de urna e o perfil demográfico da população (vide Figura A2, na qual observa-se que no Nordeste do Brasil houve uma maior concentração de urnas modelo 2020 em áreas próximas a capitais). Logo, as eventuais diferenças observadas podem ser reflexo desses perfis demográficos, bem como ideologias, perfis psicológicos, crenças e outros fatores, não do modelo de urna: seria necessária grande dose cautela para controlar essa grande diversidade de fatores em eventuais análises e, pela ausência de detalhes no Relatório Argentino, não é possível dizer se isso foi de fato feito.

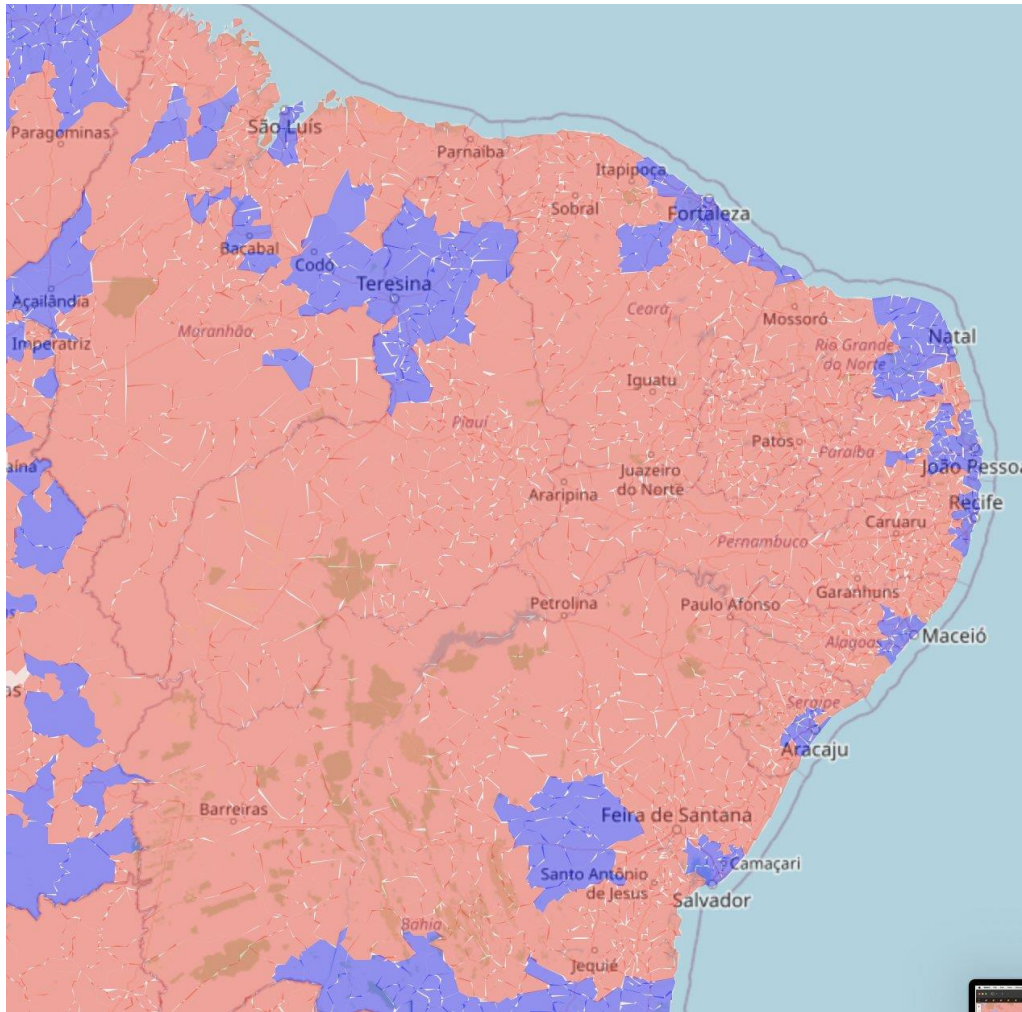
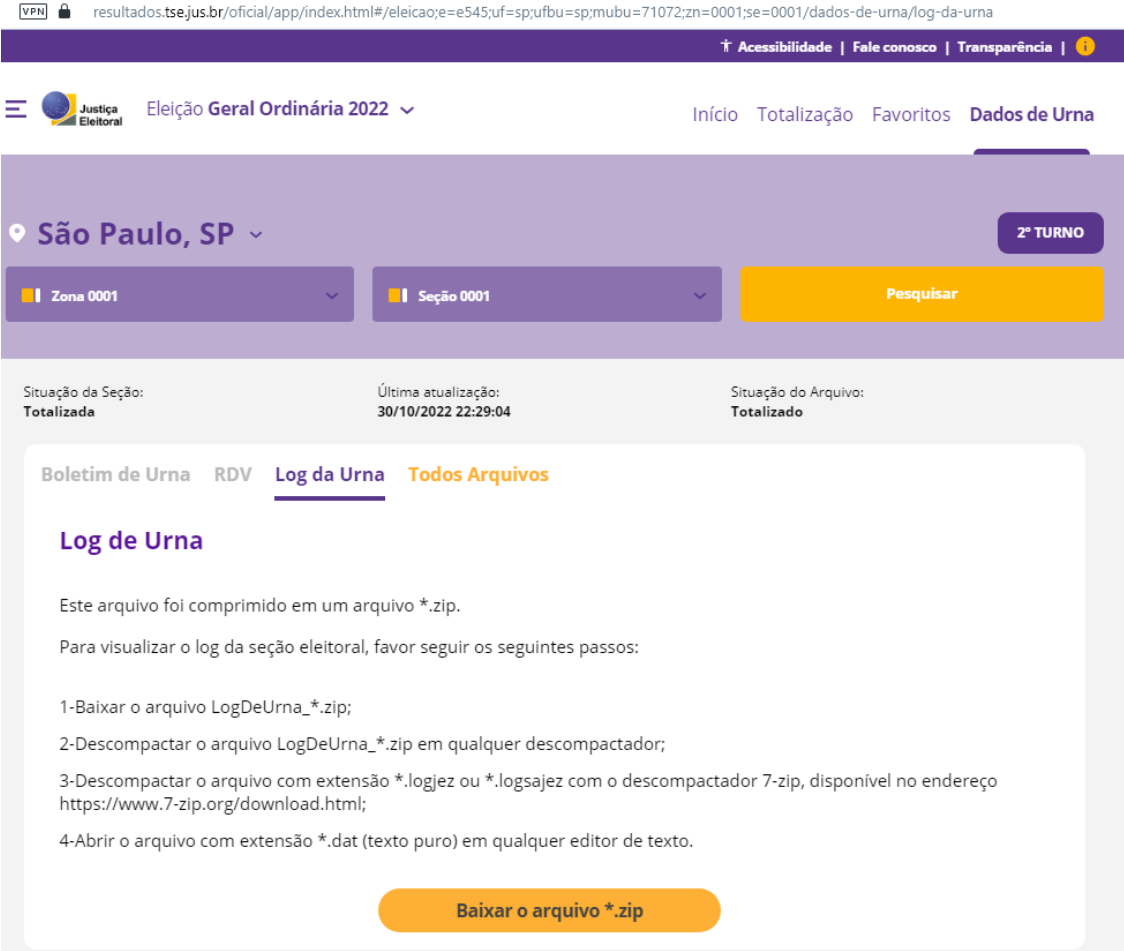


Figura A2 – Distribuição de urnas modelo 2020 (azuis) e de modelos anteriores (vermelho) na região do Nordeste. Fonte: autores, com base nos dados disponíveis em <https://dadosabertos.tse.jus.br/>.

Alegação B: Sobre diferenças em arquivos de log implicarem na existência de códigos fonte distintos

Primeiramente, cabe explicar o que significa um “log” no contexto das urnas eletrônicas. Não é nada complicado: trata-se simplesmente de um conjunto de mensagens que descrevem eventos que aconteceram na urna em questão, incluindo a data e horário de cada evento específico. Exemplos de registros desse tipo incluem o momento em que ela foi ligada, o estado da bateria então observado, o resultado das várias verificações que ela faz internamente (incluindo verificações de segurança, para identificar se o software nela carregado é legítimo), os momentos em que a votação foi iniciada e finalizada, etc.

A título de exemplo: qualquer pessoa pode acessar os arquivos de log das urnas utilizadas nas eleições de 2022, já que são dados totalmente públicos, por meio do seguinte site <https://resultados.tse.jus.br/oficial/app/index.html#/eleicao;e=e545/resultados>. Para fins de ilustração, pode-se selecionar a cidade de São Paulo/SP, Zona 001, Seção 001, e escolher a opção “Log da Urna” (vide Figura B1).



The screenshot shows the TSE website interface. At the top, there is a navigation bar with links for 'Acessibilidade', 'Fale conosco', and 'Transparência'. Below this, the main header includes the TSE logo and 'Eleição Geral Ordinária 2022'. The main content area is for 'São Paulo, SP' and '2º TURNO'. It features dropdown menus for 'Zona 0001' and 'Seção 0001', and a 'Pesquisar' button. Below these, there are three status indicators: 'Situação da Seção: Totalizada', 'Última atualização: 30/10/2022 22:29:04', and 'Situação do Arquivo: Totalizado'. The main content area has tabs for 'Boletim de Urna', 'RDV', 'Log da Urna', and 'Todos Arquivos'. The 'Log da Urna' tab is active, showing a message: 'Este arquivo foi comprimido em um arquivo *.zip. Para visualizar o log da seção eleitoral, favor seguir os seguintes passos: 1-Baixar o arquivo LogDeUrna_*.zip; 2-Descompactar o arquivo LogDeUrna_*.zip em qualquer descompactador; 3-Descompactar o arquivo com extensão *.logjez ou *.logsajez com o descompactador 7-zip, disponível no endereço https://www.7-zip.org/download.html; 4-Abrir o arquivo com extensão *.dat (texto puro) em qualquer editor de texto.' At the bottom of this section is a button labeled 'Baixar o arquivo *.zip'.

Figura B1 – Acessando o log da urna de São Paulo/SP, Zona 001, Seção 0001.

Fazendo isso, obtém-se um arquivo comprimido (extensão .zip) que, após aberto, permite acesso ao arquivo logd.dat. Esse é o arquivo de log, que nada mais é que um arquivo de texto que começa com as informações a seguir:

```

22/09/2022 11:02:54 INFO 67305985 LOGD Início das operações do logd
22/09/2022 11:02:54 INFO 67305985 LOGD Urna ligada em 22/09/2022 às 11:01:31
22/09/2022 11:02:54 INFO 67305985 SCUE Iniciando aplicação - Oficial - 1º turno
22/09/2022 11:02:54 INFO 67305985 SCUE Versão da aplicação: 8.26.0.0 - Onça-pintada
22/09/2022 11:02:56 INFO 67305985 SCUE Urna operando com bateria interna
22/09/2022 11:02:56 INFO 67305985 SCUE Bateria interna com carga parcial
...

```

O que o autor do Relatório Argentino percebeu é que, em algumas urnas, há “diferenças sutis entre os entre os arquivos de ‘log’”, mostrando um exemplo. Daí, o autor concluiu que “há, confirmadamente, ao menos dois softwares nas urnas, em diferentes modelos, não dependentes de determinado modelo”. Logo de início, a conclusão pareceu um tanto estranha, porque é muito comum que um mesmo software possa registrar eventos diferentes em seu log dependendo do que ele observa durante a execução. Um exemplo hipotético, mas ilustrativo: caso a bateria da urna esteja baixa, pode haver um alerta ao operador e um registro correspondente no log, de modo que os logs de urnas que estejam com a bateria baixa serão diferentes daqueles apresentados por urnas que não tenham esse problema durante sua inicialização. Tudo isso com **o mesmo software**.

Apesar dessa estranheza na afirmação do Relatório Argentino, decidiu-se averiguar se ela tinha algum mérito. Para isso, a equipe de pesquisadores buscou identificar, no código fonte da urna, especificamente o local onde as diferenças apontadas no relatório poderiam aparecer (as diferenças em questão são reproduzidas na Figura B2).

```

logd.dat modelo 'tipo 1'
27/09/2022 14:42:03 INFO 67305985 SCUE VerificaA_o de assinatura de dado por etapa [8] - [/dsk/fe/estatico/jez/vota_apl.vst] - [SUCESSO] 1957E6A18AB3EF65
27/09/2022 14:42:04 INFO 67305985 SCUE VerificaA_o de assinatura de dado por etapa [9] - [/dsk/fe/estatico/jez/vota_bin.vst] - [SUCESSO] 193B3FB187BD676E
27/09/2022 14:42:08 INFO 67305985 SCUE VerificaA_o de assinatura de dado por etapa [10] - [/dsk/fe/estatico/jez/vota_ofi.vst] - [SUCESSO] 89E800A7CC154A7B
27/09/2022 14:42:08 INFO 67305985 SCUE VerificaA_o de assinatura de dado por etapa [11] - [/dsk/fe/estatico/jez/votapart.vst] - [SUCESSO] 490EACCCF3B1561E
27/09/2022 14:42:21 INFO 67305985 SCUE Identifica_o do Modelo de Urna: UE2009 40E4266A08F2B03E
27/09/2022 14:42:32 INFO 67305985 SCUE Data e hora solicitada: 2A17C0A044275238
27/09/2022 14:43:04 INFO 67305985 SCUE Data e hora que foi digitada pelo operador: 27/09/2022 - 14:43:00 36AB862FB1C024D
27/09/2022 14:43:06 INFO 67305985 SCUE SeA_o informada pelo operador: 0030 1FB6F94579CE213E
27/09/2022 14:43:10 INFO 67305985 SCUE Início do particionamento da MI: 88CB4BF5388C256B
27/09/2022 14:43:12 INFO 67305985 SCUE Particionamento realizado com sucesso 23F80FB858801BFF
27/09/2022 14:43:13 INFO 67305985 SCUE Início da formataA_o 30A4600751B95860
27/09/2022 14:43:28 INFO 67305985 SCUE FormataA_o realizada com sucesso 6B7146E6EFC3EB8

logd.dat modelo 'tipo 2'
27/09/2022 19:42:36 INFO 67305985 SCUE VerificaA_o de assinatura de dado por etapa [8] - [/dsk/fe/estatico/jez/vota_apl.vst] - [SUCESSO] DAF06F4A8EEFECDF
27/09/2022 19:42:36 INFO 67305985 SCUE VerificaA_o de assinatura de dado por etapa [9] - [/dsk/fe/estatico/jez/vota_bin.vst] - [SUCESSO] DCAE996C97DDC213
27/09/2022 19:42:40 INFO 67305985 SCUE VerificaA_o de assinatura de dado por etapa [10] - [/dsk/fe/estatico/jez/vota_ofi.vst] - [SUCESSO] 2CE600170CCE3980
27/09/2022 19:42:41 INFO 67305985 SCUE VerificaA_o de assinatura de dado por etapa [11] - [/dsk/fe/estatico/jez/votapart.vst] - [SUCESSO] 7439488335576B18
27/09/2022 19:42:46 INFO 67305985 SCUE Identifica_o do Modelo de Urna: UE2009 2BF83BD2C9184B4A
27/09/2022 19:42:56 INFO 67305985 SCUE Data e hora solicitada: 4307A00300E0704E
27/09/2022 19:43:05 INFO 67305985 SCUE Data e hora que foi digitada pelo operador: 27/09/2022 - 19:34:00 D6C779864455B769
27/09/2022 19:43:25 INFO 67305985 SCUE SolicitaA_o do n_mero da seA_o 026D08D0C8A03029
27/09/2022 19:43:34 INFO 67305985 SCUE SeA_o informada pelo operador: 0212 029E3E8A885315A2
27/09/2022 19:43:39 INFO 67305985 SCUE Início do particionamento da MI: DDA98BA31700B40Z
27/09/2022 19:43:41 INFO 67305985 SCUE Particionamento realizado com sucesso 23639DA3B9AD6EAB
27/09/2022 19:43:41 INFO 67305985 SCUE Início da formataA_o 666B5CCC92CFD515
27/09/2022 19:43:56 INFO 67305985 SCUE FormataA_o realizada com sucesso 5F219EC4C84B8203

```

Figura B2 – Logs de urnas onde foram observados pontos de diferença no “Relatório argentino”

O resultado foi que, novamente, o “Relatório Argentino” está equivocado: percebeu-se que o mesmo software (a saber, o único que foi disponibilizado à USP) é capaz de gerar trechos de log contendo ou não a linha adicional apontada, dependendo de como se deu a interação com o operador da urna durante a inicialização do sistema. O vídeo do procedimento experimental consta a seguir: https://youtu.be/Lrd9YZtP_RQ.

Mais especificamente: conforme é possível verificar nos logs fornecidos, eles se dão na aplicação SCUE. O SCUE é o "Sistema de Carga da Urna Eletrônica". Esta informação pode ser obtida no site do TSE: <https://www.tse.jus.br/eleicoes/processo-eleitoral-brasileiro/logistica-e-preparacao/ecossistema-da-urna>. Este aplicativo é o responsável por carregar todo o software que roda na urna de uma mídia externa para

a sua mídia interna, o que ocorre antes da eleição, durante a chamada “carga urna”. O momento que determina se a linha adicional de log “Solicitação do número da seção” aparece ou não é apresentado na Figura B3 a seguir:

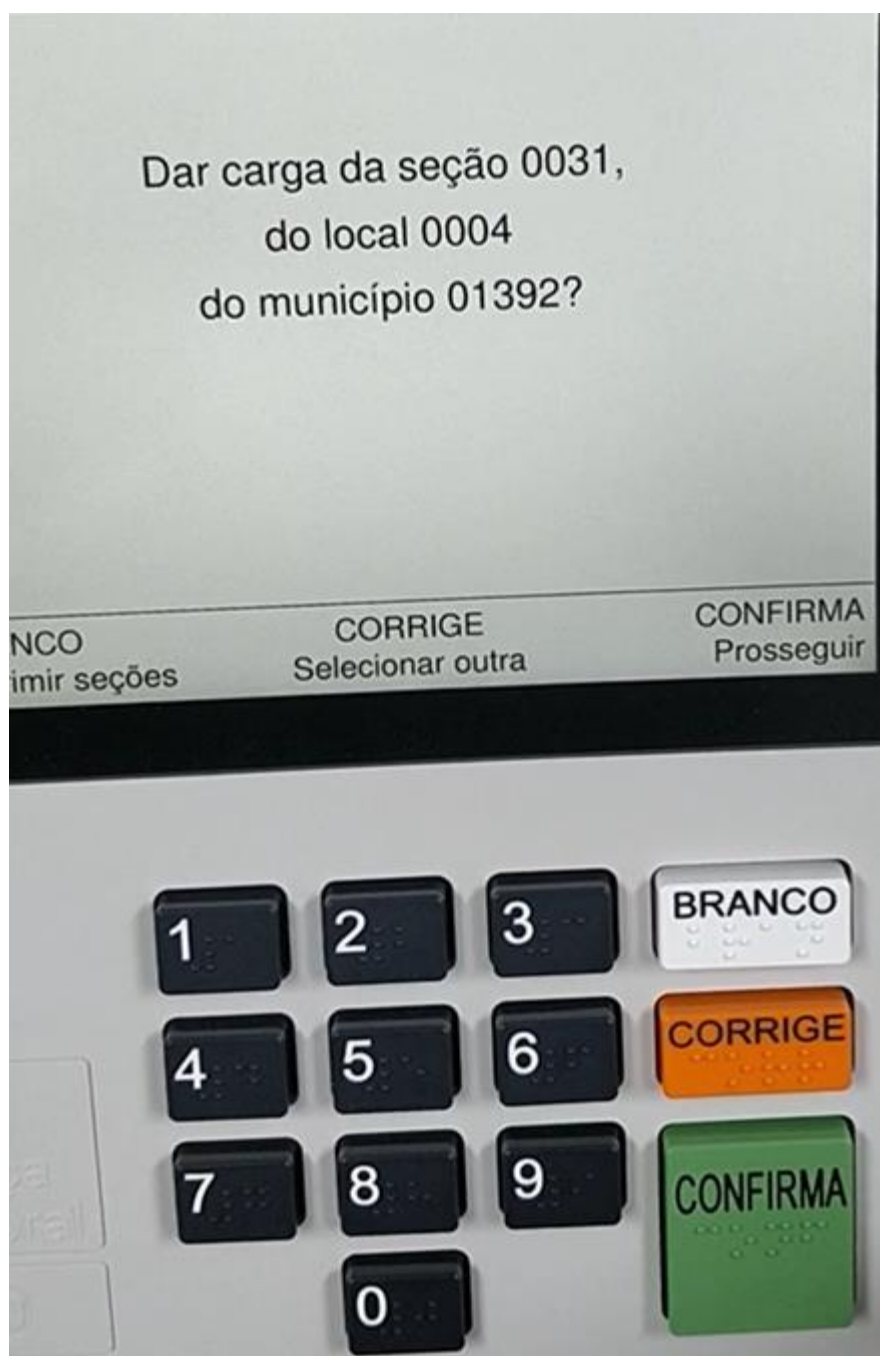


Figura B3 – Foto da urna eletrônica modelo 2020 à disposição dos investigadores na USP, mostrando o momento em que o SCUE está sendo executado.

A equipe de investigadores pôde então constatar que a linha de log aparecer ou não depende do operador da urna pressionar a tecla Confirma (assim gerando o log circulado em verde) ou pressionar a tecla Corrige (assim gerando o log circulado em vermelho). Após executar este experimento para os dois casos e extraindo os logs resultantes, foi possível reproduzir tais sequências de logs, usando o mesmo software. Dessa forma, ao contrário do que afirma o Relatório Argentino, as sequências distintas nos logs não representam indícios de existência de códigos fontes distintos. Cabe

ênfatizar que este não é um dos únicos casos da urna em que logs são introduzidos na sequência condicionalmente por fatores de interação com o usuário. Desta forma, como previamente mencionado, é aconselhada cautela em alegar indício de fraude com base em achado de sequências de logs distintas entre si.

Conclui-se, portanto, que estão totalmente equivocadas as afirmações sensacionalistas como “há, confirmadamente, ao menos dois softwares nas urnas”, que “Jamais poderia haver sequer uma diferença nesta sequência” e que “Nada mais explica essa diferença que não ao menos duas versões de softwares”. Na realidade, é muito fácil explicar essas diferenças analisando o código fonte, não havendo a priori razão alguma para assumir que a existência de códigos distintos seria a única explicação para isso (exceto, novamente, a tentativa de criar sensacionalismo).

Alegação C: Sobre a existência de uma “trava” na soma dos votos aos candidatos Lula e Bolsonaro

A afirmação da existência de uma “trava” sobre a soma do número de votos para os candidatos Lula e Bolsonaro é repetida em diversos pontos do Relatório Argentino, sendo ilustrada na Figura C1.

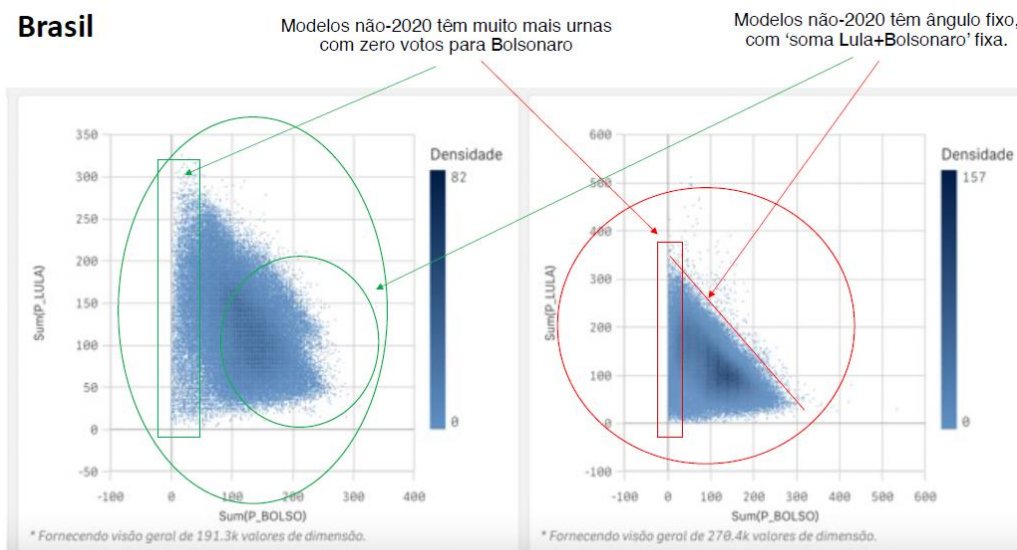


Figura C1 - Exemplo de suposta “trava” na soma de votos para Lula e Bolsonaro nos modelos não-2020 (direita), a qual não ocorreria nos modelos 2020 (esquerda). Fonte: “Relatório argentino”

O autor do relatório argumenta então que “Este jamais seria um comportamento esperado” para se referir aos gráficos das urnas de modelos anteriores à UE2020. Entretanto, esse comportamento é na realidade perfeitamente esperado, tanto para as UE2020 quanto para os modelos anteriores: simplesmente, trata-se de uma consequência natural de haver um número máximo de eleitores aptos a votar em cada urna.

Para facilitar a compreensão deste ponto, considere um cenário simplificado, no qual o número máximo de eleitores em todas as urnas seja 400, e que haja apenas 2 candidatos, o primeiro recebendo X votos e o segundo obtendo Y votos. Nesse caso, qualquer resultado da soma de $X+Y$ deve ser no máximo 400. Afinal, não pode haver mais votos do que o número de eleitores aptos a votar! Por outro lado, nem sempre a soma $X+Y$ vai dar 400, pois alguns eleitores poderiam anular seus votos, ou não comparecer ao pleito. Logo, pode-se dizer que qualquer situação em que se observa $X+Y \leq 400$ estaria dentro do esperado. Ao amostrar diversas urnas, todas elas satisfazendo essa restrição inerente (ou, se preferir, “trava matemática”), o resultado é exatamente um triângulo como o mostrado no lado direito da Figura C1. Esse cenário simplificado, e alguns pontos representando possíveis resultados válidos de votação nesse cenário, são mostrados na Figura C2.

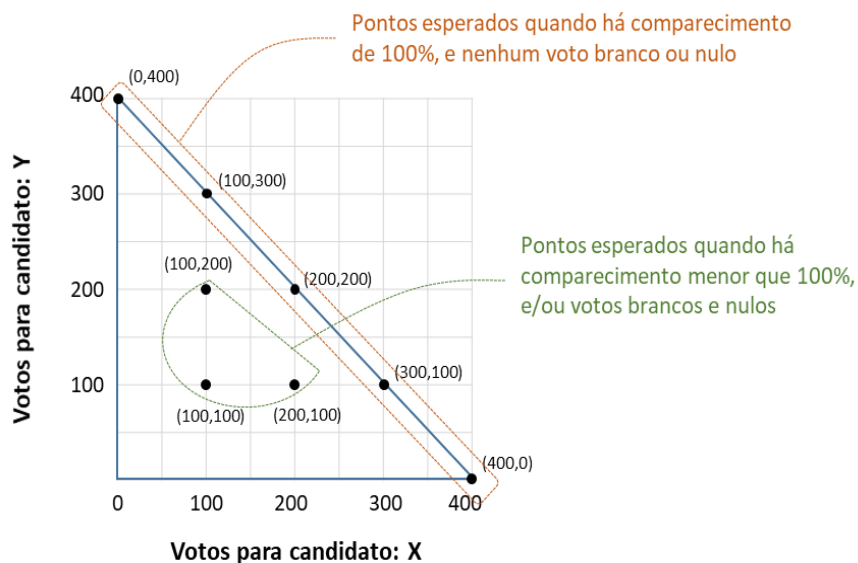


Figura C2 - Comportamento esperado para uma eleição com 2 candidatos e 400 eleitores por urna: forma-se naturalmente um gráfico de aparência triangular.

Como, mesmo no primeiro turno, os votos dos brasileiros concentraram-se essencialmente nos candidatos Lula e Bolsonaro, esse cenário simplificado com dois candidatos não é muito longe da realidade observada. Além disso, as urnas utilizadas nas eleições brasileiras costumam de fato ter um número máximo próximo de 400 eleitores, conforme mostra o gráfico da Figura C3 para as urnas efetivamente utilizadas nas eleições de 2022. A figura mostra também que há várias seções com um número menor de eleitores, embora números na faixa de 300 e 400 sejam mais comuns. Essa abordagem explica-se por que ela permite (1) reduzir o número de urnas utilizadas no pleito (mais eleitores por urna leva à necessidade de menos urnas), ao mesmo tempo que (2) evita-se a formação de longas filas (mais eleitores por urna aumenta a chance de formarem-se filas em horários de pico).

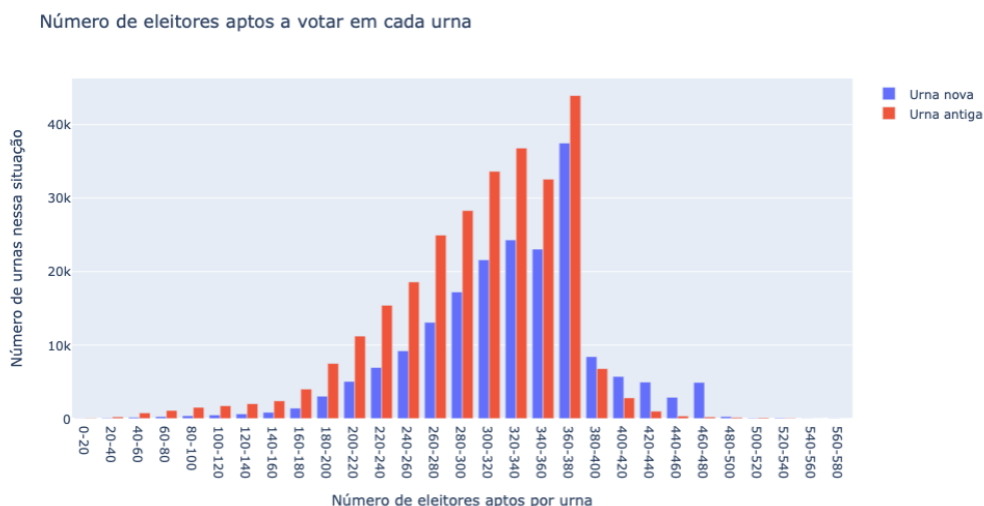


Figura C3 - Número de eleitores aptos a votar por modelo de urna, durante as eleições de 2022. Fonte: Autores com base nos dados disponíveis em <https://dadosabertos.tse.jus.br/dataset/resultados-2022/resource/9f042e8b-ed41-4b18-a01b-e0613a68d90c>

Logo, o que se esperaria observar para o gráfico criado a partir das amostragens feitas no Relatório Argentino para as UE anteriores a 2020 é exatamente o que foi observado: uma área triangular com uma zona limite concentrada na região próxima a 400!

Agora, o que explica então a “deformação” desse triângulo no gráfico da direita da Figura C1, que se refere às urnas 2020, escondendo o “ângulo fixo” observado nas urnas de modelos anteriores? Muito simples: observe novamente a Figura C3, e perceba que as UE2020 são as mais comumente usadas em seções nas quais o número de eleitores aptos supera a faixa dos 400, podendo chegar a números mais próximos de 500. Nesse cenário, espera-se que apareçam votos na região que fica após a linha que limita o número de eleitores a 400, conforme ilustra a Figura C4. Supondo que os votos nessa região mais externa se concentrem na área pontilhada mostrada à esquerda desta figura (uma área que, aproximadamente, dá um número similar de votos para os candidatos X e Y, algo estatisticamente esperado), pode-se observar algo muito similar ao perfil observado no “Relatório argentino”.

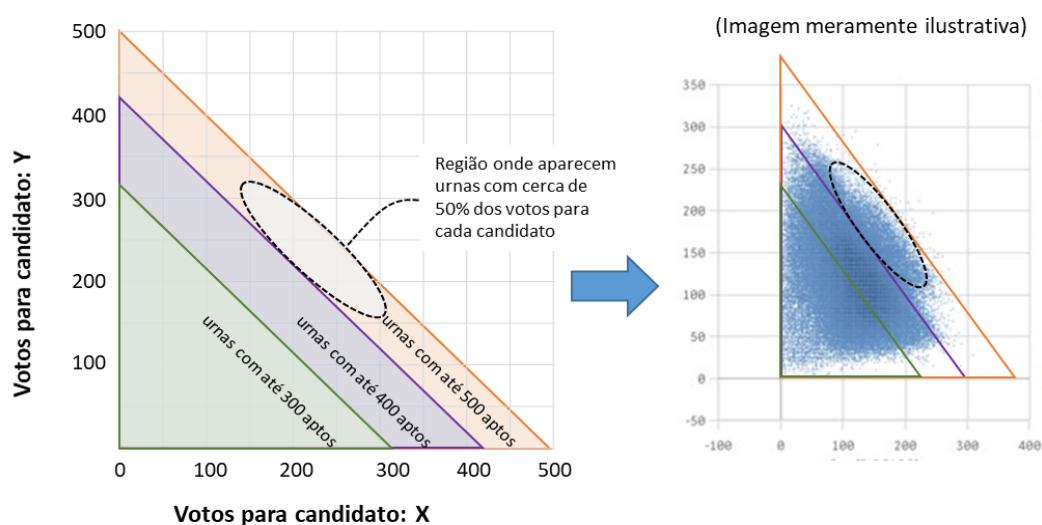


Figura C4 - Regiões onde se espera encontrar os resultados da votação para para uma eleição com 2 candidatos, para 300, 400 e 500 eleitores aptos por urna (Esquerda). Supondo que os votos da região com mais eleitores aptos se concentrem na área pontilhada, o comportamento apresentado no gráfico do Relatório argentino para as UE2020 é facilmente explicado. Nota: a imagem à direita é meramente ilustrativa, ou seja, os triângulos ali mostrados não necessariamente correspondem a urnas reais com 300, 400 ou 500 eleitores.

Novamente, ao contrário do que o Relatório Argentino afirma, não há qualquer estranheza em qualquer dos gráficos observados. Pelo contrário: usando apenas dados públicos, é fácil observar que os comportamentos obtidos, tanto para as urnas modelo 2020 como para modelos anteriores, leva exatamente a situações esperadas. Portanto, mais uma vez, o relatório equivoca-se ao fazer afirmações sensacionalistas como “Este jamais seria um comportamento esperado”.

Conclusões

Após analisar tecnicamente o assim chamado “Relatório Argentino”, as principais conclusões foram:

- **Alegação A:** alega-se não haver documentação de auditoria recente para os modelos de urna anteriores a 2020, mas apenas para estas últimas.
 - Resultado da análise: **não há qualquer embasamento** para essa afirmação, dada a existência de análises de segurança bastante recentes tanto para as urnas modelo 2020 e 2015. Mesmo considerando que modelos mais antigos foram alvo de análises similares apenas no passado, a divisão entre “urnas modelos 2020” e “urnas de modelos anteriores” carece de critério técnico e, portanto, quaisquer conclusões tiradas a partir dessa comparação não devem ter qualquer relação com o modelo em si.
- **Alegação B:** alega-se que as diferenças observadas em arquivos de log implicam na existência de códigos fonte distintos sendo utilizados nas urnas eletrônicas brasileiras.
 - Resultado da análise: embora potencialmente interessante como hipótese, **a alegação se revela falsa**. Foi possível reproduzir em laboratório o mesmo comportamento observado pelo autor do Relatório Argentino com o único código fonte da urna eletrônica recebido do TSE pela USP, sem qualquer dificuldade.
- **Alegação C:** alega-se haver uma “trava artificial” nos gráficos que mostram a soma dos votos aos candidatos Lula e Bolsonaro
 - Resultado da análise: **a alegação carece de qualquer fundamento**. Os comportamentos observados, tanto para o gráfico das urnas modelo 2020 como para os gráficos de modelos anteriores, são facilmente explicáveis pela existência de número máximo de eleitores aptos a votar em cada urna, bem como pela distribuição desses valores máximos entre as urnas do modelo 2020 e para urnas de modelos anteriores.

Dadas essas observações e explicações, o que se conclui é que **o Relatório Argentino não prima pelo rigor técnico, fazendo várias inferências infundadas e sensacionalistas. É difícil dizer se os erros (muitos dos quais podem ser considerados básicos) são causados por imperícia, pelo desejo deliberado de levar o leitor a erro, ou por uma combinação desses fatores.**

Obviamente, essas observações não reduzem o potencial interesse em se realizar uma auditoria em urnas que pareçam suspeitas aos olhos de algum candidato, caso algum deles assim o deseje. Entretanto, sugere-se fortemente que eventuais pedidos dessa natureza, caso venham a fazer uso do Relatório Argentino como parte de sua fundamentação, não se apoiem em alegações como as aqui avaliadas, pois, ao partir de premissas falsas, a lógica decorrente é que as conclusões que dali sejam obtidas também sejam errôneas.