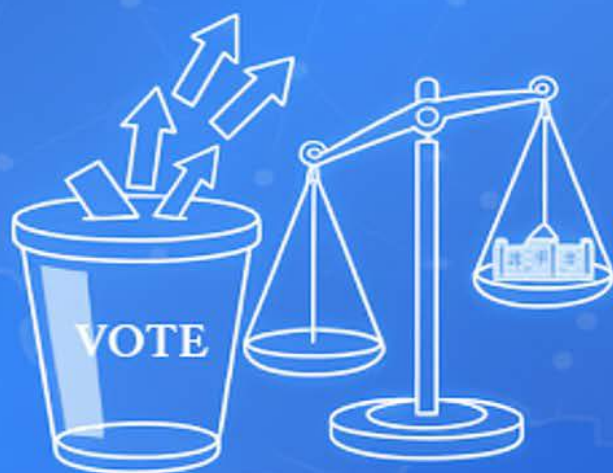


Doar un grafic să-ți mai spun

*Despre sondaje și alegeri
numai de bine*



Mircea Comșa

Presa Universitară Clujeană



Doar un grafic să-ți mai spun
Despre sondaje și alegeri numai de bine

Mircea Comșa

Presa Universitară Clujeană
2025

Referenți științifici:

Prof. univ. dr. Dan Chiribucă,
Universitatea „Babeș-Bolyai” din Cluj-Napoca

Prof. univ. dr. Bogdan Voicu,
Universitatea „Lucian Blaga” din Sibiu

ISBN: 978-606-37-2848-8

© 2025 Autorul volumului. Toate drepturile rezervate.
Reproducerea integrală sau parțială a textului, prin orice
mijloace, fără acordul autorului, este interzisă și se
pedepsește conform legii.

Universitatea Babeș-Bolyai
Presa Universitară Clujeană
Director: Codruța Săcelean
Str. B.P. Hasdeu nr. 51
400371 Cluj-Napoca, România
Tel.: (+40)-744.687.884
E-mail: editura@ubbcluj.ro
<http://www.editura.ubbcluj.ro/>
<http://libraria.ubbcluj.ro/>

Cuprins

Cuvânt înainte	8
Dinamica ratei de prezență la vot la alegerile locale din Cluj-Napoca, 1992-2024	11
Cluj-Napoca vs. Național	11
Rata reală a prezenței la vot (corecție în funcție de emigrație)	12
Cum poate fi explicată tendința de scădere a ratei de prezență la vot din Cluj-Napoca (1992-2024)?	14
Lipsa competiției electorale reale reduce prezența la vot.....	14
Eliminarea în 2011 a turului 2 al alegerilor locale	16
Alegerile locale din Cluj-Napoca -2020 vs. 2024 - variația numărului votanților și a ratei de prezență la vot în funcție de vârstă	18
Numărul votanților din Cluj-Napoca la alegerile din 2024: Locale vs. Europarlamentare.....	20
Alegerile Locale 2024: Dinamica numărului de votanți noi pe parcursul zilei.....	22
Rata de prezență la vot a tinerilor - alegerile locale și europarlamentare 2024.....	28
Cât de strânsă a fost competiția electorală?	31
Calitatea estimărilor sondajelor: alegerile Europarlamentare 2024	36
Calitatea estimărilor sondajelor: alegerile pentru funcția de primar general.....	41
AUR + Sociopol = Love?	46

Turul 2 al alegerilor prezidențiale din Moldova: o modelare a intențiilor de vot.....	50
Un oraș, doi primari	53
Turul 2 al alegerilor prezidențiale din Moldova: impactul posibil al fenomenului de cumpărare a voturilor	56
Turul 2 al alegerilor prezidențiale din Moldova: impactul posibil al ultimelor evenimente	60
Alegeri 2024 Moldova: o analiză la nivel de raion a rezultatelor oficiale	63
Alegeri 2024 Moldova: o analiză la nivel de secție de vot a rezultatelor oficiale	68
Dinamica intențiilor de vot la alegerile prezidențiale 2024 (1)...	76
„Eu cu cine votez?”	85
Dinamica intențiilor de vot la alegerile parlamentare 2024 (1) ..	86
„House effects”: alegerile prezidențiale 2024.....	94
„House effects”: alegerile parlamentare 2024	99
Alegerile prezidențiale 2024 din Moldova: transferul voturilor	103
Dinamica intențiilor de vot: alegerile generale 2024	106
Alegerile generale 2024: „house effects”	111
Alegerile generale 2024: F2F vs. CATI.....	114
Prezidențiale 2024: intențiile de vot înainte și după decizia CCR.....	118
Lupta din adâncurile sondajelor.....	128
Alegerile prezidențiale 2024, turul 1: o estimare a intențiilor de vot	132
Călin file de poveste.....	140
Update estimări T1.....	142

Prezidențiale 2024, T1: dinamica ratei de prezență.....	146
Prezidențiale 2024, T1: rata de participare în funcție de vârstă	148
Un altfel de grafic.....	151
Parlamentare 2024: o estimare a intențiilor de vot	153
... tot Parlamentare, ceva mai fancy.....	157
Numărul votanților în funcție de vârstă / anul nașterii	159
Parlamentare 2024: dinamica ratei de prezență	163
Rata de prezență la vot în funcție de anul nașterii	165
Dacă „sondaje nu e, nimic nu e!” Sau	167
Prezidențiale 2024, T2: o estimare a intențiilor de vot.....	169
Opinie: De ce candidați independenți? De ce doi?	173
De ce candidați independenți?	173
De ce doi candidați?	174
La final, o „MICĂ” problemă!	176
Calitatea estimărilor sondajelor: Prezidențiale 2024, T1	178
Estimările sondajelor realizate în primele trei săptămâni de campanie.....	179
Estimările sondajelor realizate în ultima săptămână de campanie.....	180
Medii și modelări ale estimărilor din sondaje	181
Estimările sondajelor din ziua votului	182
Sub/Supra-estimări în ziua votului	183
Despre transparența sondajelor	185
Calitatea estimărilor sondajelor: Parlamentare 2024	187
Estimările sondajelor realizate în primele trei săptămâni de campanie.....	188
Estimările sondajelor realizate în ultima săptămână de campanie.....	188
Medii și modelări ale estimărilor din sondaje	189
Estimările sondajelor din ziua votului	190
Sub/Supra-estimări în ziua votului	191

Who does (not) vote in Romania?	194
O estimare a transferului de voturi dinspre partide spre candidați	196
„Da’ liberalii un’ se duc când se duc?”	202
Profile electorale: Prezidențiale 2024, T2	206
Prezidențiale 2025: cote pariuri (1)	217
Prezidențiale 2025: cote pariuri (2).....	219
Clickbait sau lipsă de informare?	221
Prezidențiale 2025: cote pariuri (3).....	223
O estimare într-un context electoral marcat de incertitudini...	225
O estimare tot la nivel agregat	228
„House effects”: alegerile prezidențiale 2025, T1 (1).....	230
Dinamica intențiilor de vot: turul 1 al alegerilor prezidențiale din mai 2025 (1)	233
Cât de „disciplinat” e electoratul partidelor?	237
Relația dintre încredere și vot.....	241
Dinamica intențiilor de vot: turul 1 al alegerilor prezidențiale din mai 2025 (2).....	251
„House effects”: alegerile prezidențiale 2025, T1 (2)	257
Efecte de metodă: CATI vs. Online	262
Cine e pe locul 2?	266
Variația estimărilor intențiilor de vot în funcție de partizanatul potențial al institutelor de sondare	269
Dinamica intențiilor de vot: turul 1 al alegerilor prezidențiale din mai 2025 (3).....	272
Au schimbat dezbatările șansele percepute ale candidaților? ...	278

Alegerile prezidențiale 2025, turul 1: o predicție a intențiilor de vot	282
Prezidențiale 2025, T1: dinamica ratei de prezență.....	286
Prezidențiale 2025, T1: rata de prezență la vot în funcție de vârstă.....	288
Calitatea estimărilor sondajelor: Prezidențiale 2025, T1	290
Estimările sondajelor realizate în ultimele 10 zile	291
Predicțiile pe baza estimărilor din sondajele publice	293
Medii și modelări ale estimărilor din sondaje	294
Estimările sondajelor din ziua votului	295
Sub/Supra-estimări în ziua votului	296
Câteva concluzii.....	299
Despre transparența sondajelor	300
Prezidențiale 2025, T2: dinamica T2 anterior T1	302
Prezidențiale 2025: dinamica T2.....	305
Alegerile prezidențiale 2025, turul 2: o predicție a intențiilor de vot	309
Prezidențiale 2025, T2: dinamica ratei de prezență	313
Prezidențiale 2025, T2: rata de prezență la vot în funcție de vârstă.....	315
Calitatea estimărilor sondajelor: Prezidențiale 2025, T2	317
Estimările sondajelor realizate în ultimele 7 zile	318
Medii și modelări ale estimărilor din sondaje	319
Estimările sondajelor din ziua votului	320
Sub/Supra-estimări în ziua votului	321
Despre transparența sondajelor	323
Două grafice cât zece mii de cuvinte.....	324
Drumul de la T1 la T2.....	326
Land doesn't vote	336
Schimbarea preferințelor de vot la nivel de UAT	341

Cuvânt înainte

În vara anului 2024, imediat după alegerile locale și europarlamentare din România, am început blogul [Doar un grafic să-ți mai spun](#) postând inițial rar, cam de două-trei ori pe lună, apoi tot mai des, ajungând în campanie la două-trei analize pe săptămână. Un an și ceva mai târziu aveam deja 78 de postări pe teme legate de alegeri. La sugestia mai multor prieteni, fără a fi pe deplin convins că merită acest demers, am decis totuși să le adun într-un volum. Le mulțumesc celor care au avut răbdarea și promptitudinea de a citi unele dintre postări înainte de publicare și, mai ales, pentru sugestiile care le-au îmbunătățit. Le sunt recunoscător și celor care au urmărit textele de pe blog. Fără interesul și reacțiile lor, motivația de a scrie ar fi fost, probabil, mult mai mică, iar multe dintre aceste analize ar fi rămas doar pe calculatorul meu.

Anii electorali 2024 și 2025 au reprezentat un maraton politic pentru România, culminând în cazul alegerilor prezidențiale cu rezultate surprinzătoare și reconfigurări neașteptate ale scenei politice. De la alegerile locale și europarlamentare, trecând prin cele parlamentare și ajungând la un scrutin prezidențial anulat și reluat, dezbaterea publică a fost dominată de opinii, speculații și predicții adesea contradictorii. În acest peisaj informațional aglomerat și adesea confuz, prin analizele de pe blog, am încercat să ofer repere cât mai obiective, menite să-i ajute pe cei interesați să distingă semnalul de zgomot (desigur, cu referire la sondajele de opinie). Fiecare postare are în centru una sau două întrebări principale la care răspund pe baza analizelor statistice a datelor disponibile. În spatele întrebărilor din postări se află probabil două întrebări majore: (1) Care este dinamica reală a intențiilor de vot dincolo de titlurile de moment și sondajele individuale? și (2)

Cum putem descâlci cât de cât îțele atât de complicate ale competițiilor electorale?

Răspunsul pe care îl propun în această colecție de texte se bazează pe analiza riguroasă a datelor. Sub titlul sugestiv „Doar un grafic să-ți mai spun. Despre sondaje și alegeri numai de bine”, cartea de față reunește articolele pe teme electorale publicate pe blogul meu. Lucrarea nu oferă opinii politice, ci instrumente de înțelegere, transformând cifrele brute și adesea aride în povești inteligibile despre alegeri și, mai ales, despre alegătorii din România (câteva analize sunt despre alegerile din Moldova). În acest context, fiecare grafic devine un punct de plecare pentru o discuție nuanțată, o invitație la reflecție critică asupra realităților politice și informațiilor vehiculate în presă.

Volumul explorează o serie de teme esențiale ale studiilor electorale. O primă direcție majoră este analiza participării la vot. Cititorul va descoperi dinamica prezenței la urne în Cluj-Napoca în comparație cu media națională, va înțelege cum factori precum emigrația pot distorsiona cifrele oficiale și va explora cauzele scăderii participării, de la lipsa unei competiții electorale reale până la modificări legislative cu impact direct asupra rezultatelor. Sunt analizate în detaliu profilurile demografice ale votanților, ritmul în care aceștia se prezintă la urne pe parcursul zilei și particularitățile votului în rândul tinerilor.

O secțiune consistentă a cărții este dedicată unui subiect de maxim interes și controversă: calitatea sondajelor de opinie. Volumul prezintă o serie de analize ale performanței companiilor de sondare a opiniei publice în diversele competiții electorale, de la Primăria Capitalei la alegerile europarlamentare și prezidențiale. Sunt explicate și cuantificate „efectele de institut” (*house effects*) – tendința sistematică a unor institute de a favoriza anumiți candidați – și este analizat impactul metodologiei (sondaje telefonice vs. online) asupra rezultatelor. Și, nu în ultimul rând, cartea propune și compară o serie de modele de predicție a rezultatelor alegerilor.

Dincolo de cifrele brute, analizele realizate urmăresc să deslușească modul în care se reconfigurează preferințele electoratului. Prin tehnici de inferență ecologică, sunt estimate

matricile de transfer a voturilor, arătând, de exemplu, cum s-au redistribuit voturile între partide și candidați la alegerile generale sau cum s-a reconfigurat sprijinul electoral între turul I și turul II al alegerilor prezidențiale. De asemenea, sunt schițate profile ale electoratelor principalilor candidați, corelând opțiunea de vot cu atitudini față de democrație, economie sau perioada comunistă.

„Doar un grafic să-ți mai spun” este, în esență, un manual de utilizare a datelor electorale. Cartea se dorește a fi o resursă valoroasă pentru jurnaliști, analiști politici, studenți și pentru orice cetățean dornic să navigheze informat prin complexitatea democrației românești. Vă invit să parcurgeți textele nu pentru a găsi răspunsuri definitive, ci pentru a învăța să puneți întrebări, să identificați datele utile, să faceți analizele potrivite și să prezentați rezultatele într-o formă accesibilă. Vă invit să descoperiți poveștile ascunse în spatele fiecărui grafic din acest volum.

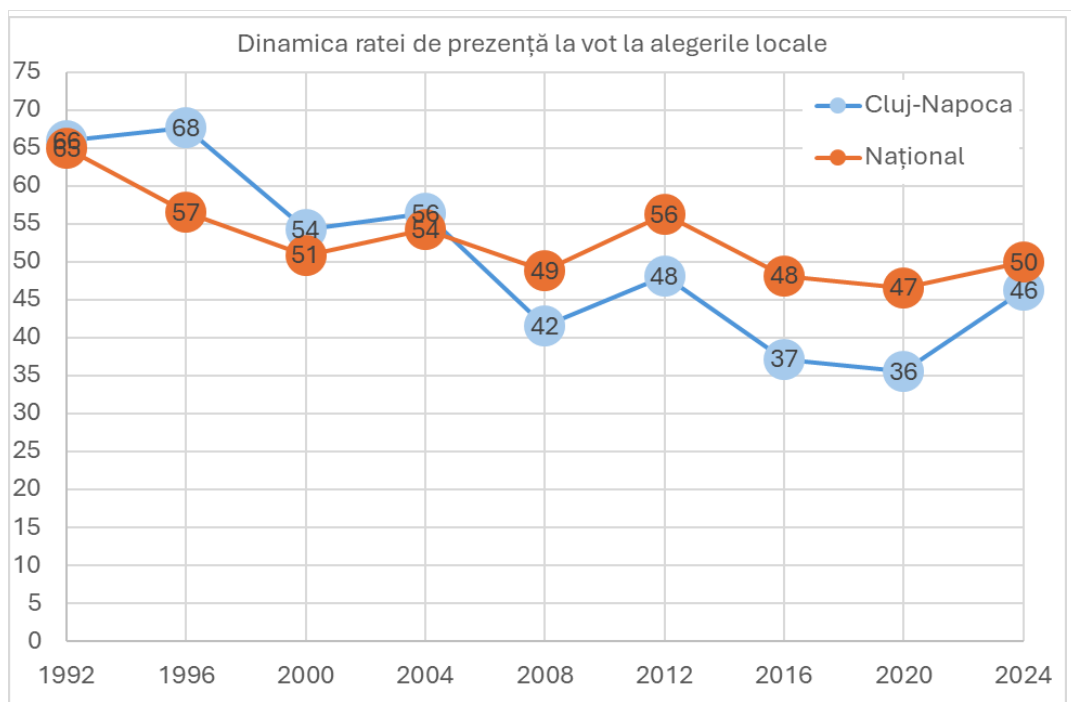
Dinamica ratei de prezență la vot la alegerile locale din Cluj-Napoca, 1992-2024

Data: 2024-06-18

Cuvinte cheie: alegeri-locale, rata-de-prezenta-la-vot

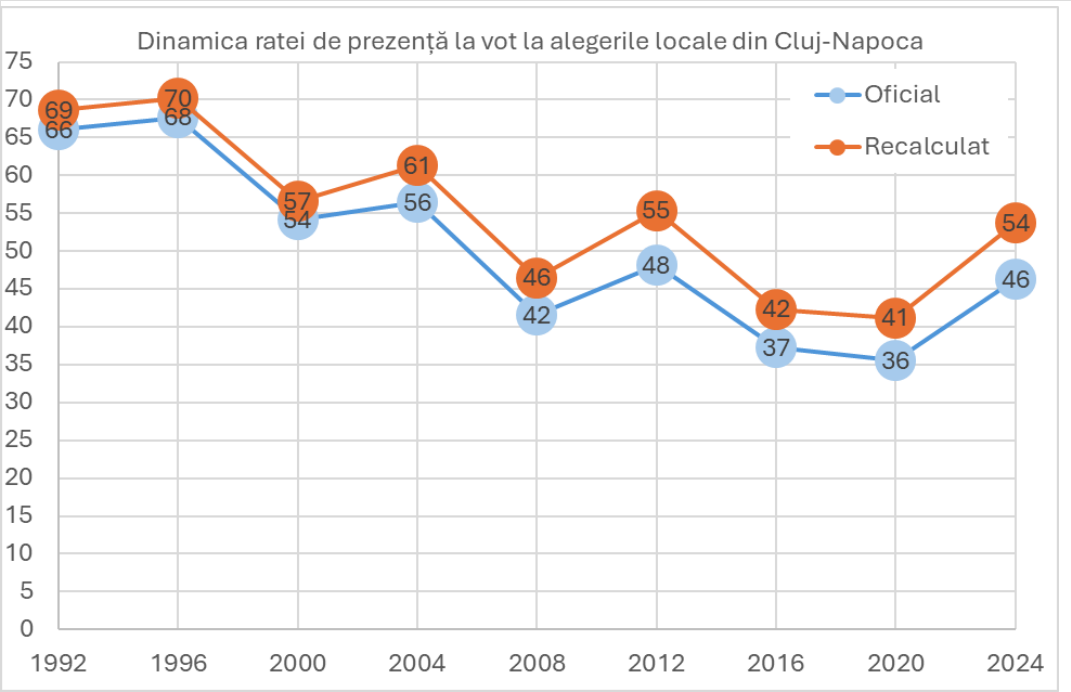
Cluj-Napoca vs. Național

La nivel național rata prezenței la vot a scăzut până în 1996, apoi a fluctuat în zona 47-57%. În Cluj-Napoca rata de prezență la vot a scăzut până în 2024, deși se observă tendințe de creștere temporară în cazul unora dintre alegerile locale. În 2024 rata de prezență la vot a crescut mult comparativ cu 2020 și a ajuns aproape să o egaleze pe cea din 2012. Până în 2004 rata de prezență la vot a fost mai mare în Cluj-Napoca comparativ cu media la nivel național, ulterior a fost mai mică. În 2024 rata de prezență la vot în Cluj-Napoca s-a apropiat de media națională.



Rata reală a prezenței la vot (corecție în funcție de emigrație)

La alegerile din perioada 1992-2024 se constată o tendință de scădere a ratei de prezență la vot în Cluj-Napoca indiferent de modul de calcul. Rata reală de prezență la vot este ceva mai mare ca urmare a efectului emigrației: cei plecați din țară nu pot vota la alegerile locale în secțiile de vot din străinătate. Deoarece ponderea emigranților a crescut în timp (după COVID19 a scăzut treptat), **scăderea reală a ratei de prezență la vot este ceva mai mică decât o indică cifrele oficiale.**



3

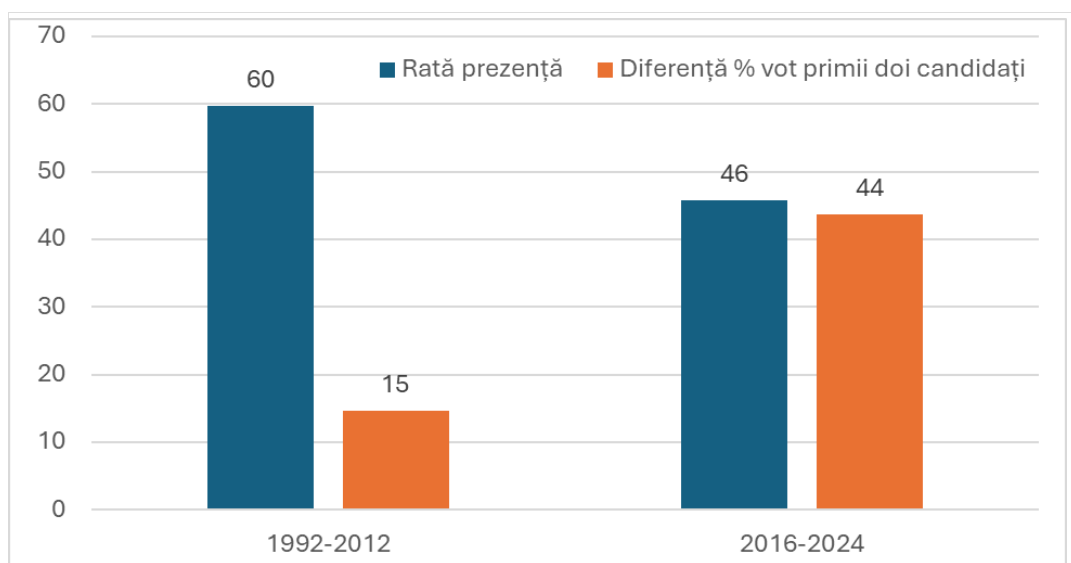
Cum poate fi explicată tendința de scădere a ratei de prezență la vot din Cluj-Napoca (1992-2024)?

Data: 2024-06-19

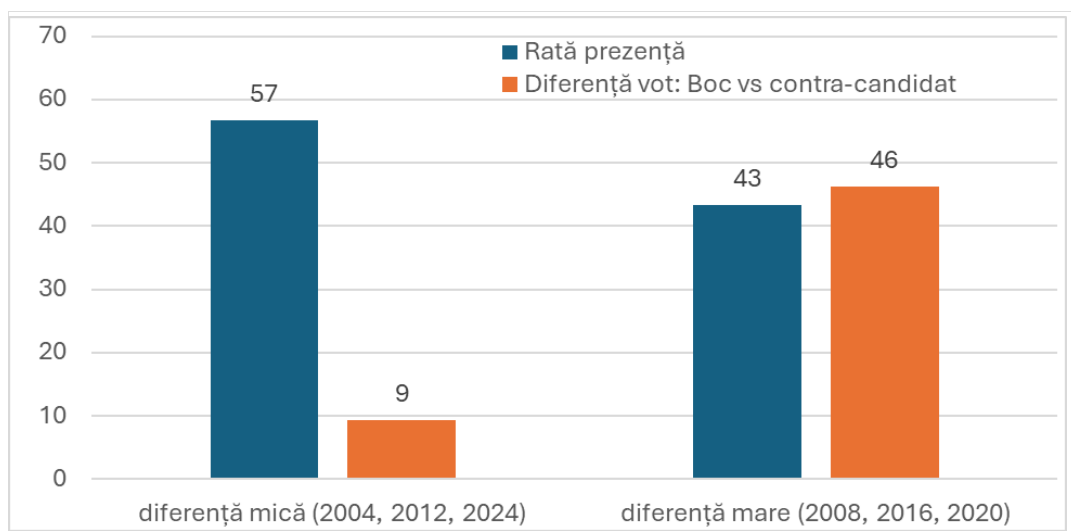
Cuvinte cheie: rata-de-prezenta-la-vot

Lipsa competiției electorale reale reduce prezența la vot

În Cluj-Napoca, în general, primarul în funcție a condus în competițiile electorale. Atunci când diferența dintre primii doi candidați a fost mai mică, rata de prezență la vot a fost relativ mai mare. Atunci când diferența dintre primii doi candidați a fost mare, prezența la vot a fost relativ mai mică. Prin urmare, o competiție redusă aduce cu sine o prezență mai mică la vot.



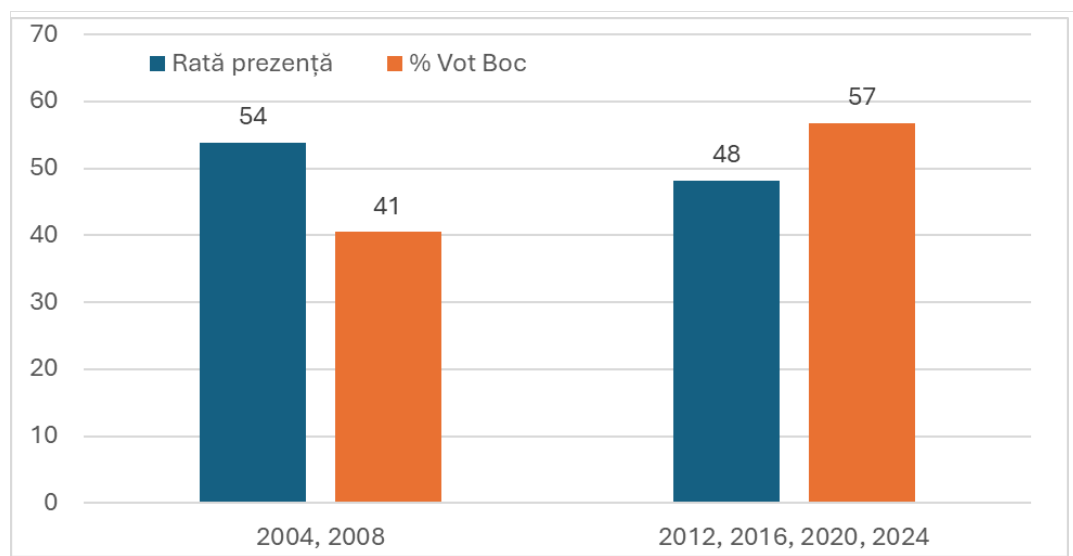
Efectul de cauzalitate semnalat anterior se aplică și în cazul actualului primar. La alegerile pentru primar din 2004, 2012 și 2024 rata de prezență a fost mai mare, iar diferența de vot dintre Emil Boc și principalul contra-candidat a fost mai mică. În ceilalți ani rata de prezență a fost mai mică, iar diferența dintre Emil Boc și principalul contra-candidat a fost mai mare. Prin urmare, o competiție redusă favorizează primarul în funcție.



Eliminarea în 2011 a turului 2 al alegerilor locale

În 2011 (Guvernul Boc 2), la propunerea unor parlamentari PDL (în prezent cei mai mulți sunt în PNL), legea electorală pentru alegerile locale a fost modificată în sensul trecerii la votul într-un singur tur. Anterior, primarii erau aleși în două tururi. Dacă la primul tur al alegerilor numărul de voturi primite de candidatul cu cel mai mare număr de voturi nu depășea jumătate plus unu din numărul total al alegătorilor înregistrați, se organiza turul doi între primii doi candidați ca număr de voturi.

La alegerile pentru primar anterioare lui 2011 (2004 și 2008) rata de prezență a fost mai mare, iar ponderea votului pro-Boc a fost mai mică. La alegerile din 2012 și următoarele rata de prezență a fost mai mică iar ponderea votului pro-Boc a fost mai mare. Prin urmare, ca urmare a acestei schimbări legislative, Emil Boc și alți primari aflați într-o situație similară au avut șanse mai mari să câștige alegerile ulterioare.





Alegerile locale din Cluj-Napoca - 2020 vs. 2024 - variația numărului votanților și a ratei de prezență la vot în funcție de vârstă

Data: 2024-06-20

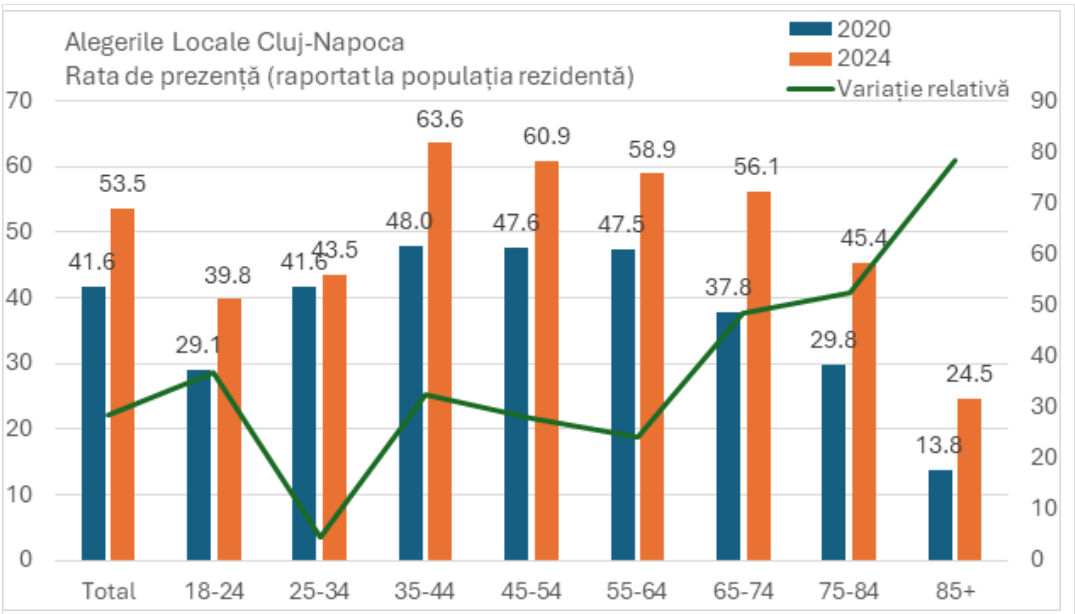
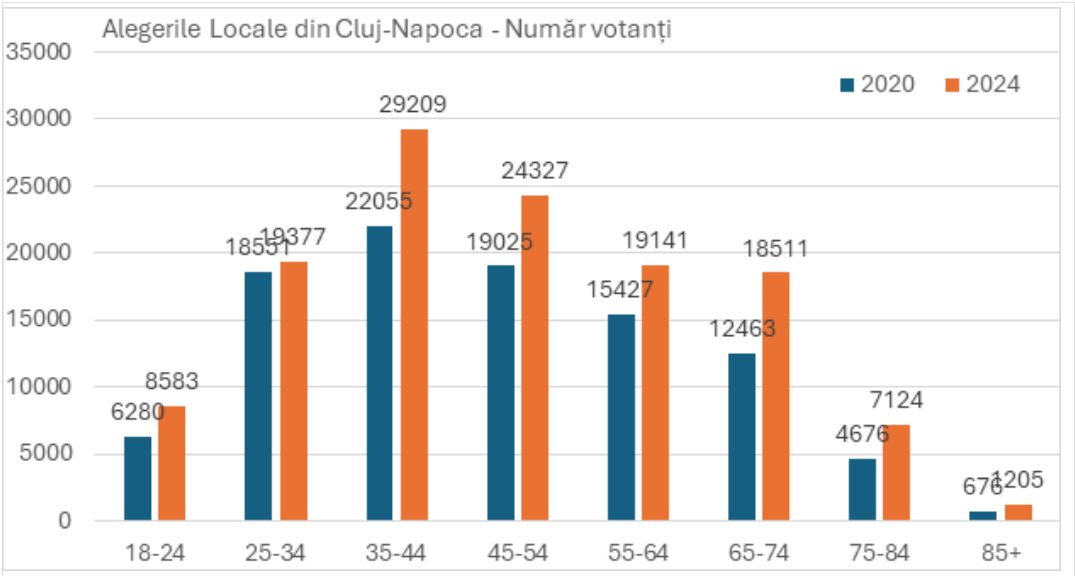
Cuvinte cheie: alegeri-locale, rata-de-prezenta-la-vot

La alegerile locale din acest an, comparativ cu 2020, în Cluj-Napoca numărul de votanți a crescut cu aproximativ 28 mii. S-a observat o creștere a numărului de votanți în rândul tuturor categoriilor de vârstă. Creșterea relativă a fost mai mare în cazul seniorilor – aceștia au revenit la vot după retragerea observată în cazul alegerilor locale din 2020 (ca urmare a pandemiei COVID19).

Foarte probabil modificarea opțiunilor de vot nu a fost produsă de schimbarea structurii pe vârste a votanților (poate cel mult marginal). Au venit votanți noi, din toate categoriile de vârstă, cel mai probabil nemulțumiți de starea actuală a lucrurilor din Cluj-Napoca sau care doresc o schimbare a situației actuale și care au identificat prezența în competiția electorală a unui contracandidat cu șanse de succes (și cu o campanie electorală atractivă). Desigur, mobilizarea realizată de către competitorii electorali a avut și ea o contribuție importantă (semnificativ mai mare comparativ cu alegerile anterioare).

Rata de prezență la vot a fost ușor mai mare în cazul bărbaților. Totuși, ponderea bărbaților în total votanți a fost puțin mai mică (e ceva și mai mică în total alegători).

Rata de prezență la vot a crescut cu vârsta până în zona 35-64 ani, apoi a scăzut. Tinerii (18-24) au o rată de prezență la vot similară cu seniorii (75+).



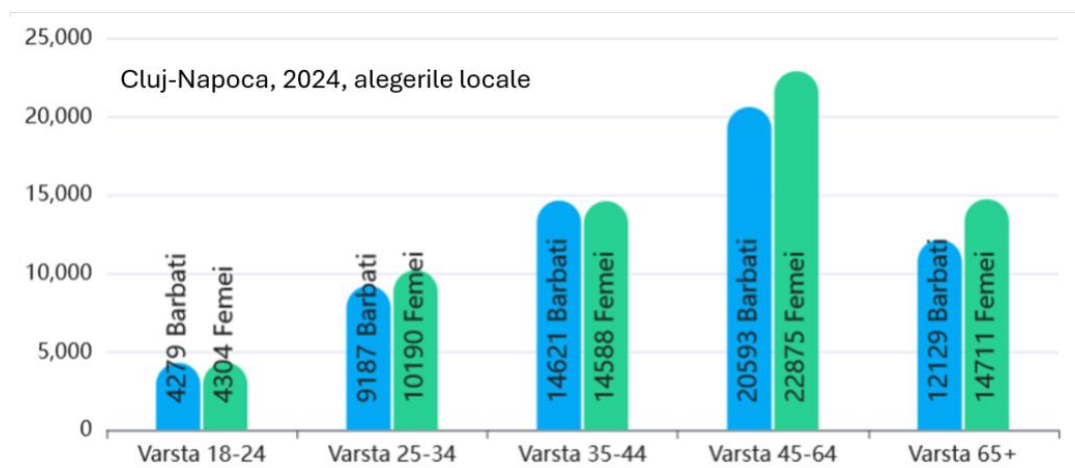
5

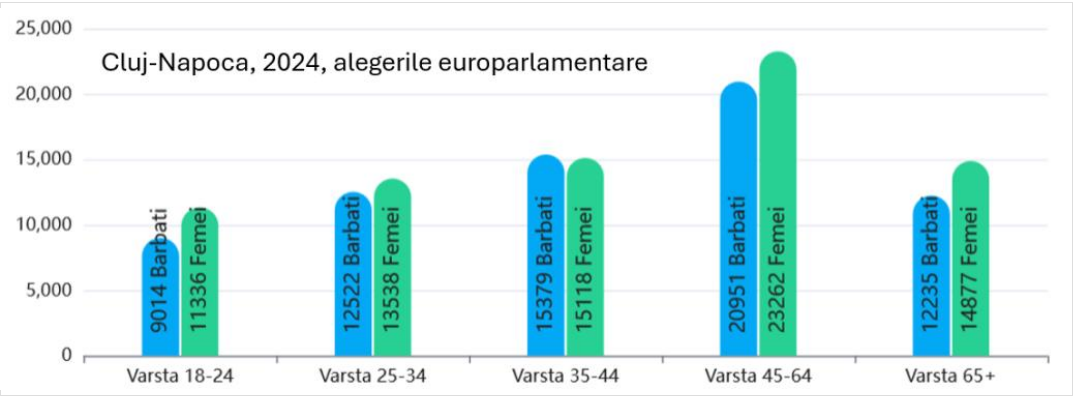
Numărul votanților din Cluj-Napoca la alegerile din 2024: Locale vs. Europarlamentare

Data: 2024-06-21

Cuvinte cheie: alegerile-europarlamentare-2024

În 2024 au fost aproximativ 20.000 voturi în plus la alegerile europarlamentare din Cluj-Napoca, majoritatea acestora aparținând foarte probabil studenților. De aici și dublarea numărului de voturi asociate tinerilor (18-24) în cazul alegerilor europarlamentare comparativ cu cele locale. La restul categoriilor de vârstă creșterile au fost mult mai mici. Această diferență de structură a votanților (vârstă și educație) explică parțial diferențele dintre rezultatele obținute de competitori la cele două tipuri de alegeri.





Alegerile Locale 2024: Dinamica numărului de votanți noi pe parcursul zilei

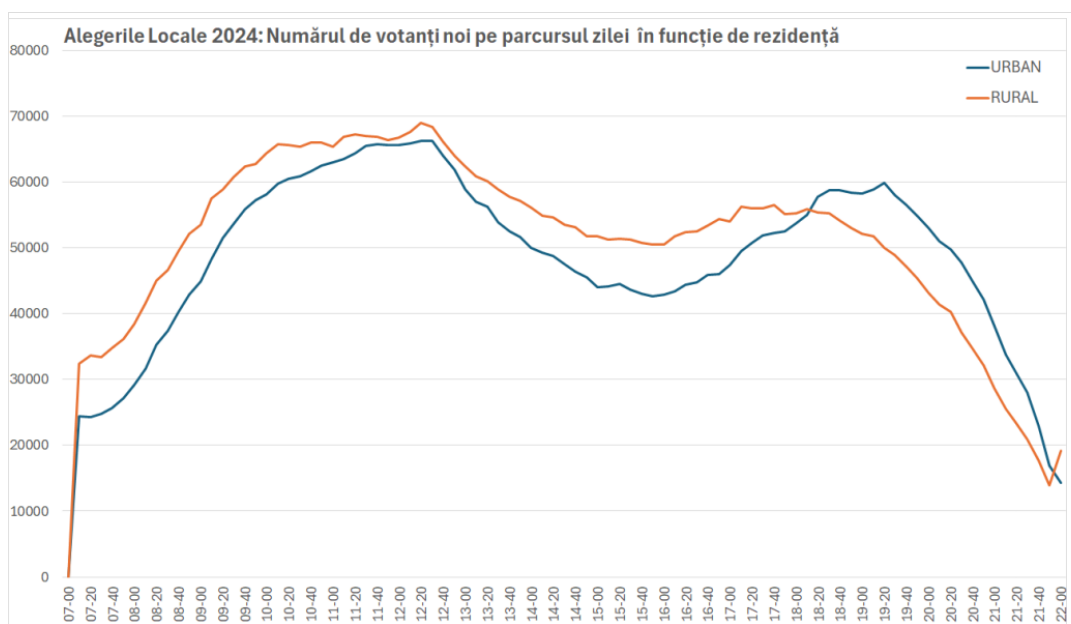
Data: 2024-06-22

Cuvinte cheie: alegeri-locale

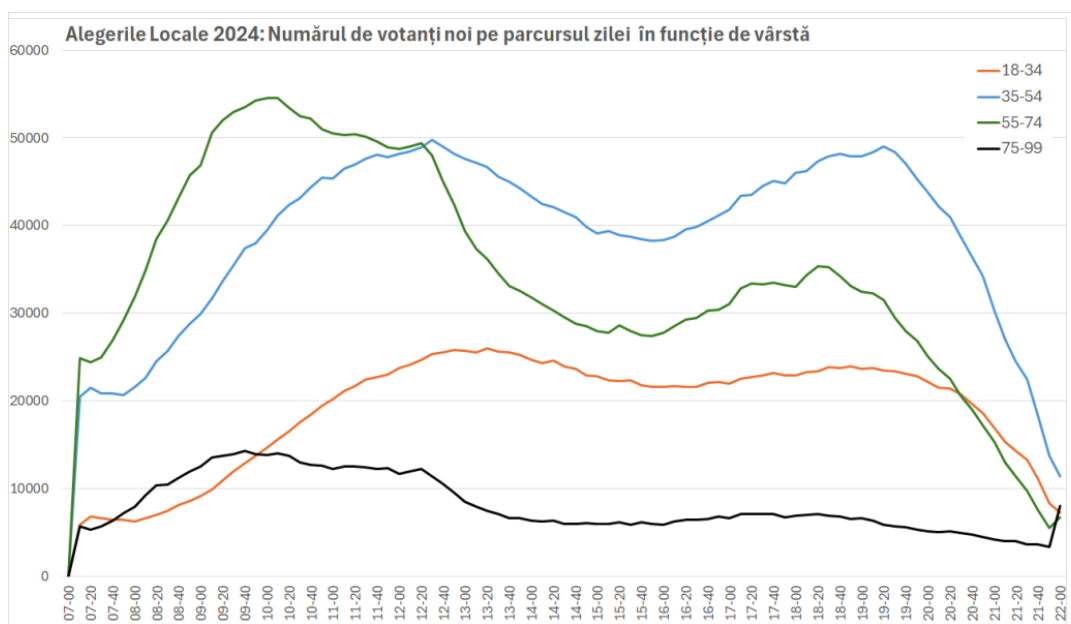
Datele cu privire la variația numărului de votanți pe parcursul zilei de vot arată că există un patern al ritmului votării. Fiecare rundă de alegeri urmează acest patern cu unele variații determinate de tipul alegerilor, perioada din an în care sunt organizate alegerile și particularitățile competiției (candidați, evenimente semnificative, incertitudinea rezultatului final).

Alegerile locale din 2024 au urmat în mare măsură paternul general. În prima jumătate de oră s-a observat o creștere abruptă a prezentării la urne, apoi ritmul votării a încetinit puțin dar a rămas foarte ridicat până în jurul orei 10. Între 10 și 12:30 observăm o fază de platou în zona superioară a prezenței, numărul celor care votează fiind mare și relativ constant. În continuare numărul votanților scade până în jurul orei 16, după care urmează al doilea val. Acesta nu mai ajunge la înălțimea primului val și atinge maximul local în jurul orei 19. Ulterior, numărul votanților scade relativ constant până la închiderea urnelor.

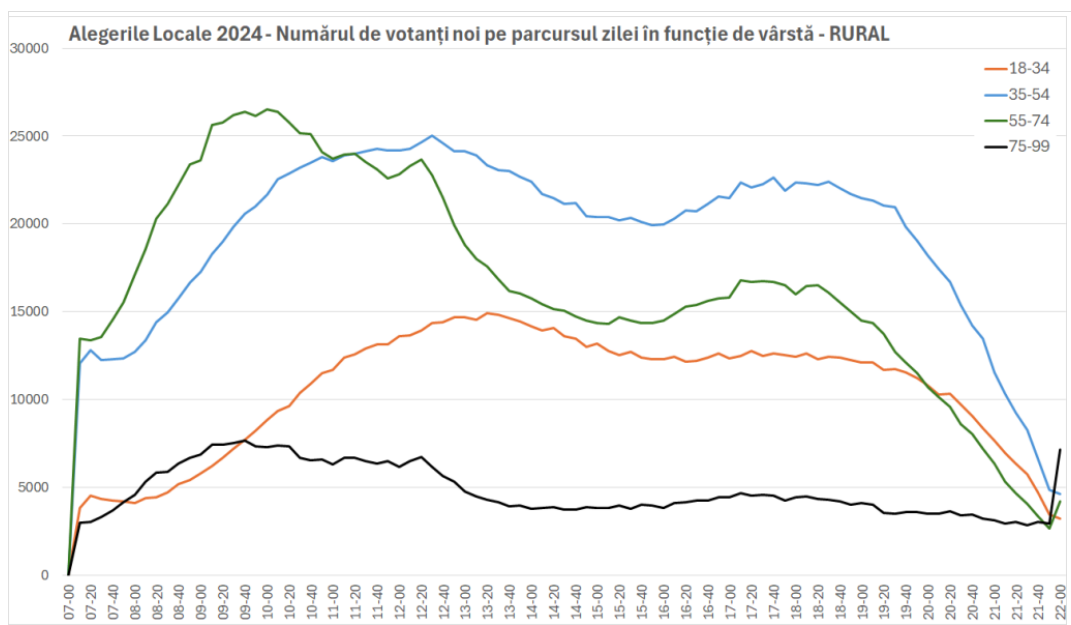
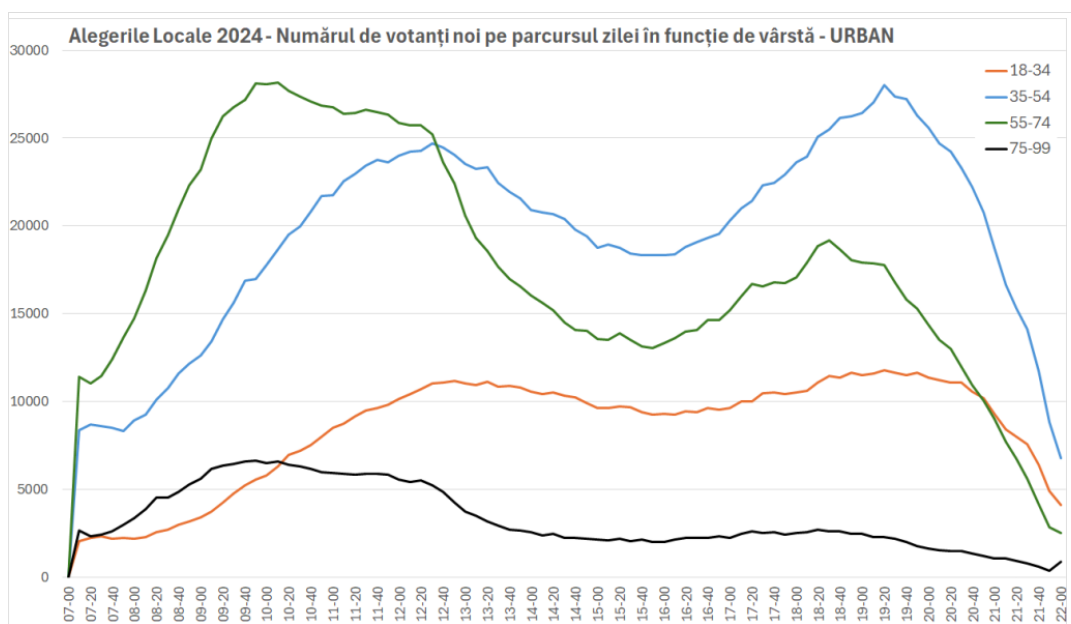
Ritmul votării nu diferă fundamental în funcție de mediul de rezidență: ruralul conduce de la început până spre prima parte a serii, după care urbanul preia conducerea.



Ritmul votării diferă semnificativ mai mult în funcție de vârstă. Seniorii (75+) votează mai degrabă constant pe tot parcursul zilei, fiind totuși ceva mai mulți dimineața. Votanții din intervalul de vârstă 55-74 ani sunt cei mai matinali; ulterior, intensitatea votării în cazul acestei categorii scade deși se constată o oarecare creștere spre ora 18. Cei de 35-54 ani preiau conducerea după ora 12:40 și domină votarea până la sfârșit, având un a doua vârf de prezență în jurul orei 19:20. Numărul tinerilor votanți (18-34) crește lent până în jurul amiezii, apoi intră într-o zonă de relativ platou până spre seară, după care scade.



Ritmul votării în funcție de vârstă nu diferă prea mult între urban și rural. Cu privire la urban, categoria de vârstă 35-54, observăm că valul doi este ceva mai înalt și mai scurt. O altă diferență dintre urban și rural apare în cazul celor cu vârstă 75+ (de fapt 65+). Observăm că în rural există un număr foarte mare de voturi chiar înainte de închiderea urnelor (ultimele 10 minute), majoritatea acestora fiind ale unor persoane în vârstă (65+). O astfel de situație este mai degrabă contra-intuitivă, neașteptată și face notă discordantă cu restul datelor. De aici și întrebarea cu privire la care sunt factorii posibili care au dus la această creștere. Ne putem imagina două ipoteze principale. Acestea sunt mai degrabă complementare, nu concurente.



Prima ipoteză se referă la momentul înregistrării în SIMPV a votanților care au folosit urna mobilă. Conform recomandării BEC, cererile pentru urna mobilă se depun în preziua alegerilor (intervalul 18-19) sau în ziua alegerilor (anterior orei 14) ([Hotărârea BEC 65/2024](#)). După aprobarea cererilor, CNP-urile

solicitanților sunt preînregistrate în SIMPV. Ulterior, doi reprezentanți ai secției de votare însoțiți de un reprezentant al MAI se deplasează la domiciliul solicitanților având cu ei toate materialele necesare votării. După ce urna mobilă este adusă în secția de vot, obligatoriu înainte de ora 22, CNP-urile sunt înregistrate definitiv în SIMPV. Dat fiind faptul că urna mobilă poate fi adusă în secția de vot oricând până la ora 22, e posibil ca, în unele cazuri, CNP-urile asociate voturilor din urnă să fi fost înregistrate definitiv în SIMPV doar la final, în ultimele 10 minute, și nu treptat.

O a doua ipoteză se referă la existența unui proces de mobilizare a alegătorilor de către candidați, mai ales spre finalul zilei, ca urmare a anticipării unui rezultat mai puțin favorabil. Presa și observatorii au oferit de-a lungul timpului evidențe diverse în acest sens. Cum s-a petrecut concret o astfel de mobilizare și cât de legale au fost măsurile întreprinse e dificil de spus strict în baza datelor disponibile. Ne putem imagina scenarii de tipul „aducere cu arcanul la oaste” (asigurarea transportului la secția de vot), convingerea și ajutarea alegătorilor să voteze folosind urna mobilă sau chiar folosirea cărților de identitate ale acestora pentru a vota. Faptul că în cazul alegerilor locale, chiar mai mult comparativ cu restul alegerilor, mai ales în rural, se folosesc recompense și amenințări pentru a crește numărul voturilor favorabile, respectiv pentru a-l reduce pe cel al voturilor nefavorabile nu e deloc o noutate („Zăhărelul” și „nuiaua.” [O incursiune în lumea practicilor electorale ilegale](#)). Incidența mai mare a acestui tip de situații în mediul rural are legătură cu faptul că monitorizarea prezenței la vot a fiecărui alegător e mai ușor de realizat, costă semnificativ mai puțin, iar recompensa relativă asociată fiecărui vot este semnificativ mai mare.

Efectul semnalat apare mai pregnant doar în cazul unora dintre localitățile din rural. Există un număr mic de localități în care numărul votanților în vârstă din ultimele 10 minute a crescut brusc. Mai exact, în 77% dintre localitățile rurale nu a fost nicio creștere, în 15% din cazuri creșterea a fost mai mică de un punct procentual (pp), în 4% creșterea a fost de 1<2 pp, iar în restul de 4% din cazuri creșterea a fost de 2 pp sau mai mult (raportat la totalul votanților). Probabilitatea ca un procent din totalul

votanților să se prezinte voluntar la vot în ultimele 10 minute e destul de mică. Probabilitatea ca majoritatea acestora să fie și în vârstă e și mai mică. Situația observată mai ales în cazul ultimei categorii de localități poate ridica o serie de semne de întrebare cu privire la corectitudinea alegerilor. Desigur, dacă nu e cumva un caz de înregistrare a voturilor din urna mobilă doar la final.

Dintre cele 124 localități rurale în care s-a înregistrat o pondere foarte mare a votanților în vârstă în ultimele 10 minute, primele două sunt Zâmbreasca (Teleorman) și Smulți (Galați). În aceste localități ponderea din total alegători a votanților în vârstă care au fost înregistrați în SIMPV în ultimele 10 minute a fost de 9%, respectiv 10%.

În fiecare dintre cele două localități sunt puțin peste 1000 alegători din care ceva mai mult de 700 au votat la alegerile locale (deci rata de prezență a fost în jur de 70%). Numărul voturilor cu urna mobilă a fost 111 / 77, iar numărul voturilor pe liste speciale a fost 6 / 44. Din datele oficiale rezultă că în ultimele 10 minute au votat 68/71 persoane cu vârsta de cel puțin 65 ani (la restul categoriilor de vârstă apar, în total, mai puțin de 10 votanți în ultimele 10 minute). Dat fiind faptul că numărul voturilor cu urna mobilă a fost mare, destul de probabil voturile înregistrate în ultimele 10 minute au fost de fapt voturile celor care au votat cu ajutorul urnei mobile pe parcursul zilei.

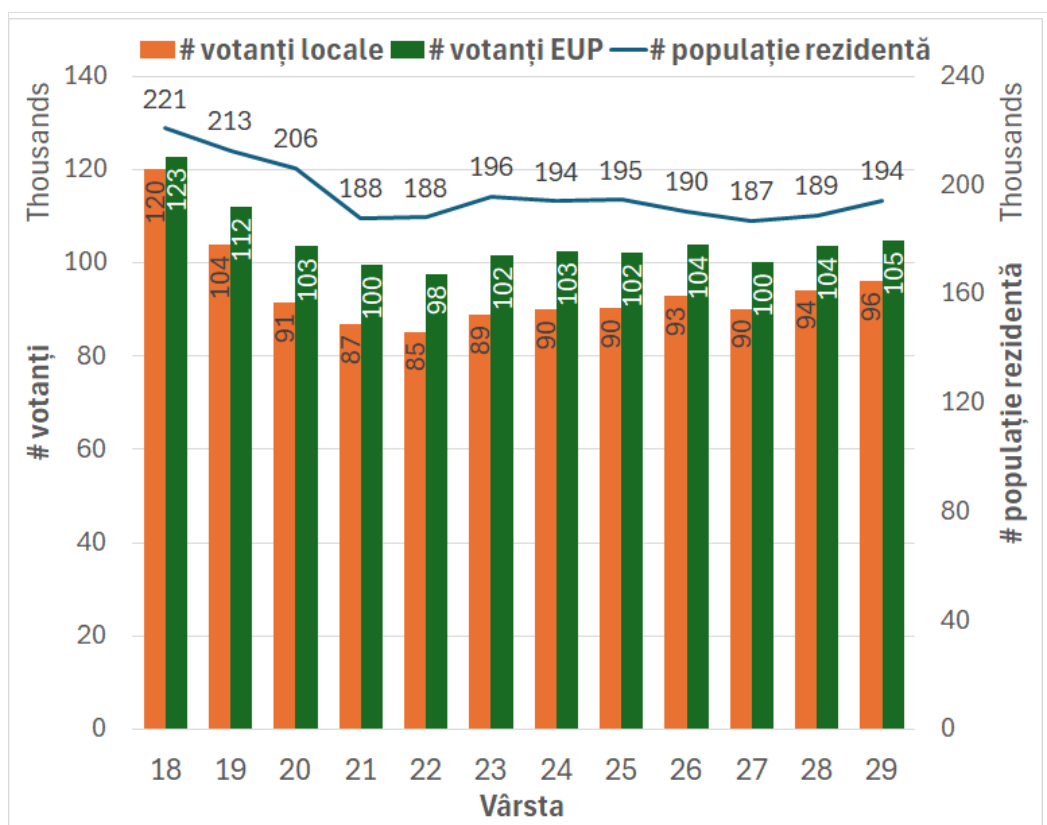
În ambele localități au fost doi competitori principali (la Smulți a fost și un candidat din partea PRO România), primarul în funcție din partea PSD și un contra-candidat din partea PNL. În ambele cazuri primarul PSD în funcție a câștigat un nou mandat. Diferența dintre primii doi candidați a fost de 67, respectiv 94 voturi.

Rata de prezență la vot a tinerilor – alegerile locale și europarlamentare 2024

Data: 2024-06-23

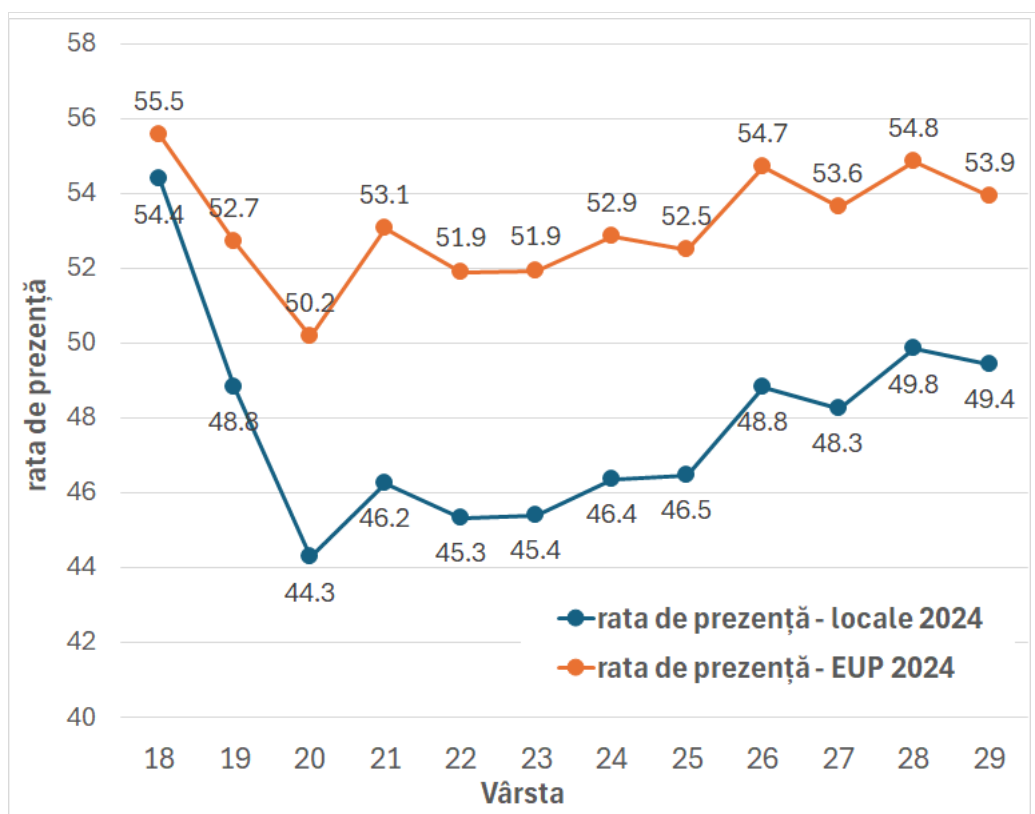
Cuvinte cheie: alegerile-europarlamentare-2024, rata-de-prezenta-la-vot

Fiecare dintre cohortele anuale ale celor care în prezent au 18-29 este compusă din aproximativ 190-220 mii persoane rezidente. În funcție de vârstă și tipul alegerilor, aproximativ 85-120 mii dintre aceste persoane au votat la alegerile locale sau europarlamentare din acest an. Cele mai recente trei grupuri (18, 19 și 20 ani) includ ceva mai multe persoane, iar primele două grupuri (18 și 19 ani) includ și un număr puțin mai mare de votanți.



Rata de prezență la vot variază în funcție de vârsta tinerilor. Prezența este mai ridicată la 18 ani, scade până la 20, apoi crește treptat până ajunge aproape de cea observată la punctul de plecare (18 ani). Acest patrn apare în cazul ambelor tipuri de alegeri (locale și europarlamentare) cu o diferență importantă: scăderea este mult mai mare în cazul alegerilor locale. De ce are loc această scădere și de ce apare o diferență între cele două tipuri de alegeri?

Cea mai mare parte a tinerilor de 18-19 ani locuiesc cu părinții. Ulterior acestei vârste, o parte a tinerilor se mută în altă localitate pentru studii. În general părinții tinerilor participă la vot într-o măsură mai mare. Cât timp tinerii locuiesc cu părinții șansele de transmitere, voluntară sau nu, a acestui comportament dezirabil social sunt mai mari. Prin urmare, o parte mai mare a tinerilor participă la vot. Dat fiind faptul că la alegerile locale se poate vota doar în localitatea de domiciliu (sau dacă alegătorul are viză de flotan), scăderea ratei de prezență la vot este mai mare în cazul acestor alegeri.



Dacă tot ne plângem de „dezinteresul” tinerilor pentru alegeri, probabil nu ar strica să ținem cont și de contextul diferit în care se află (o parte dintre) aceștia. Cum am putea face asta? Două măsuri posibile sunt scăderea vârstei de vot la 16 ani și introducerea votului online. **Reducerea vârstei de la care se poate vota ar oferi tinerilor doi ani suplimentari de socializare într-un cadru relativ mai favorabil actului de exercitare a dreptului de vot. Votul online ar reduce semnificativ costul asociat votării în cazul celor plecați la studii. Și nu în ultimul rând, ar face votul „cool.”** Foarte probabil efectul ar apărea și în cazul altor tipuri de alegători. În plus, nu am mai avea nevoie de două săptămâni pentru a afla rezultatul alegerilor, am evita renumărarea voturilor, oboseala și stresul celor implicați în proces, discuțiile și contestațiile. Desigur, funcție și de cum e implementată, o astfel de modificare ar putea produce alte probleme. Însă, putem beneficia de experiența unor țări precum Estonia.

8

Cât de strânsă a fost competiția electorală?

Data: 2024-06-25

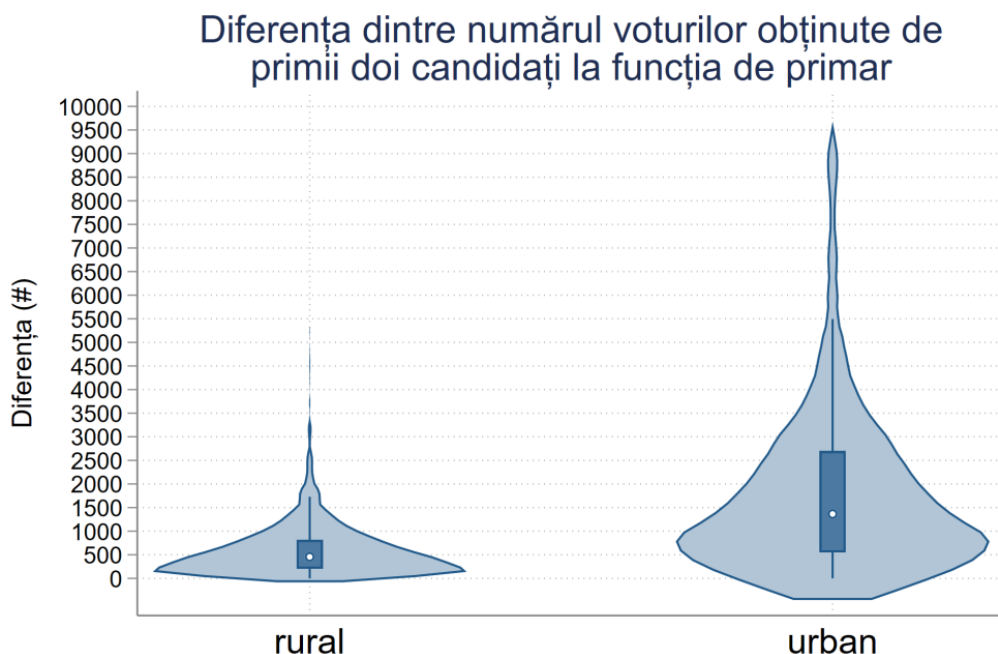
Cuvinte cheie: alegeri-primar

Rezultatul alegerilor locale definește în mare măsură configurația în care va fi realizată guvernarea localității în noul ciclu electoral, adică determină care vor fi obiectivele urmărite de administrația locală, cum vor fi alocate resursele, respectiv ce măsuri vor fi luate la nivel local. Uneori competițiile electorale sunt câștigate detașat de unul dintre candidați. Într-un astfel de caz vorbim de o competiție slabă. Alteori, diferența dintre primii doi (sau chiar trei) candidați este foarte mică, deci vorbim de o competiție strânsă.

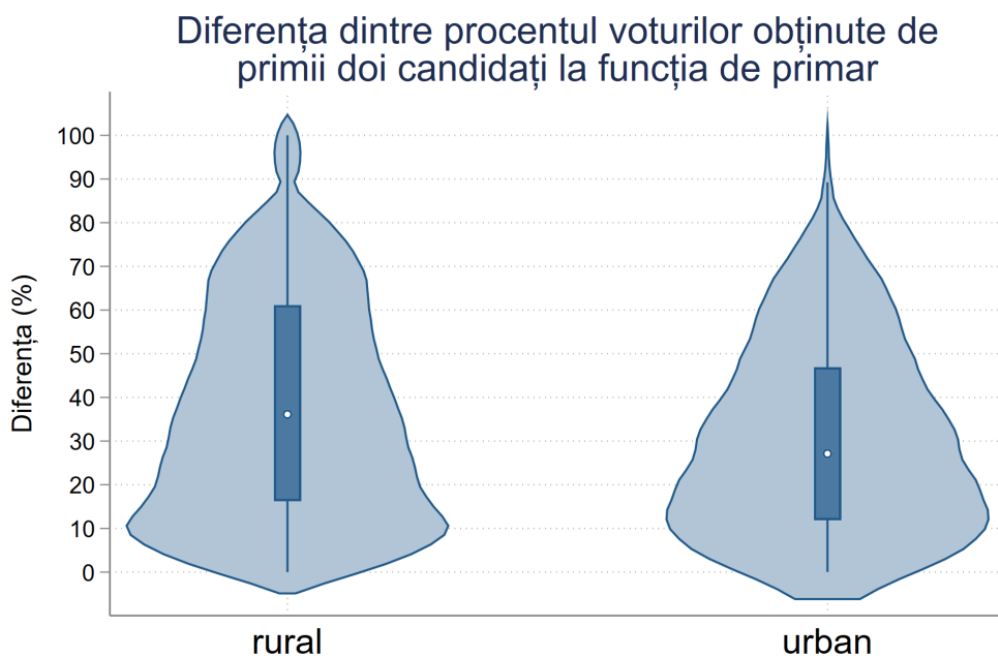
În această postare analizez global competițiile pentru funcția de primar din localitățile României în cazul alegerilor locale din 2024. Îi am în vedere doar pe primii doi clasăți și măsoz distanța dintre ei în două moduri: diferența dintre numărul de voturi obținute (diferența absolută) și diferența dintre numărul de procente obținute (diferența relativă).

Dacă ne uităm la numărul voturilor obținute, competițiile din comune au fost semnificativ mai strânse comparativ cu cele din orașe. Astfel, **jumătate din competițiile** din comune s-au decis la o diferență mai mică de 456 voturi în timp ce în orașe diferența a fost de 1486 voturi. Similar, un **sfert din competițiile** din comune s-au decis la o diferență mai mică de 205 voturi în timp ce în orașe diferența a fost de 582 voturi. Dat fiind faptul că în orașe locuiesc

semnificativ mai mulți alegători decât în comune, această situație era de așteptat.

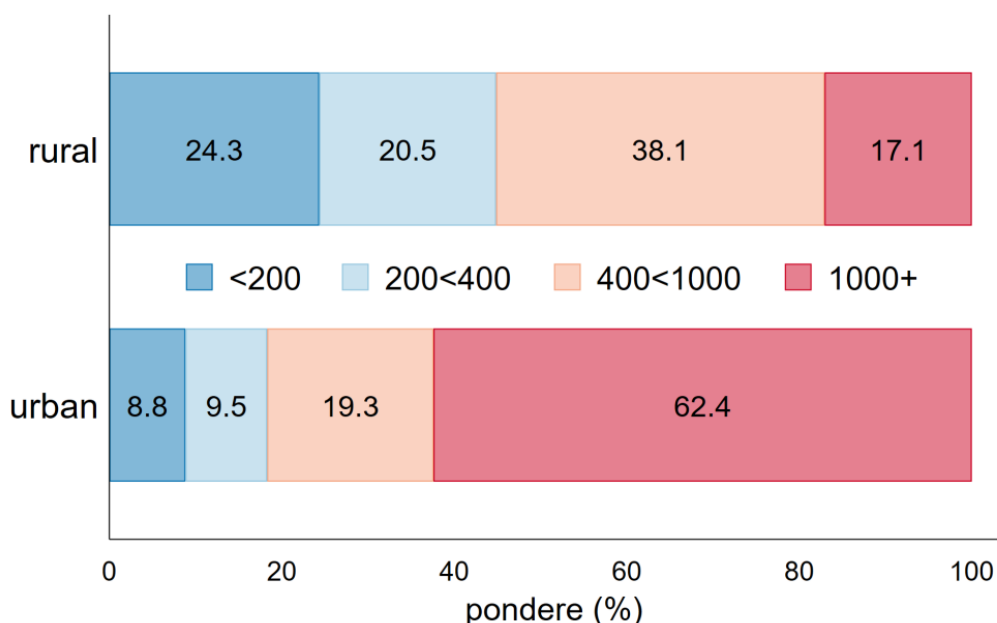


Dacă ne uităm la diferența dintre procentele voturilor obținute ajungem la o concluzie diferită față de cea descrisă mai sus: în acest caz, deși diferențele sunt mai degrabă mici, competiția este mai strânsă în urban comparativ cu rural. Altfel spus, în orașe, comparativ cu în comune, candidatul clasat pe prima poziție câștigă alegerile la o distanță mai mică față de candidatul clasat pe poziția secundă. Astfel, **jumătate din competițiile** din orașe s-au decis la o diferență mai mică de 27 puncte procentuale (pp) în timp ce în comune diferența a fost de 36pp. Similar, un **sfert din competițiile** din orașe s-au decis la o diferență mai mică de 12pp în timp ce în comune diferența a fost de 16pp.



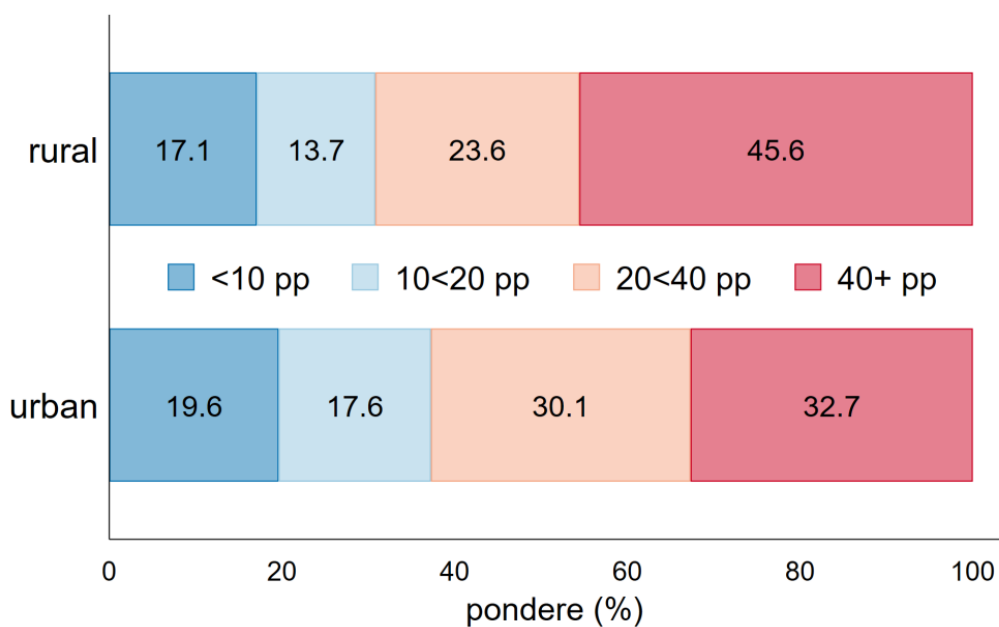
Comparația dintre urban și rural poate fi realizată și în funcție de anumite praguri comune. În acest caz am grupat diferența dintre numărul de voturi obținute de primii doi candidați în patru categorii: sub 200, $200 < 400$, $400 < 1000$ și 1000 voturi sau mai multe. Observăm că o diferență mai mică de 200 voturi apare în cazul a 9% dintre localitățile urbane, respectiv 24% dintre localitățile rurale. La polul opus, o diferență de 1000 voturi sau mai multe se regăsește în cazul a 62% dintre orașe, respectiv 17% dintre comune.

Ponderea localităților în funcție de diferența dintre numărul voturilor obținute de primii doi candidați la funcția de primar



Diferența dintre procentele obținute de primii doi candidați este grupată tot în patru categorii: sub 10 pp (puncte procentuale), 10<20 pp, 20<40 pp și 40+ pp. O diferență mai mică de 10 pp voturi apare în cazul a 20% dintre localitățile urbane, respectiv 17% dintre localitățile rurale. La polul opus, o diferență de 40 pp sau mai mare apare în cazul a 33% dintre orașe, respectiv 46% dintre comune. Practic, având în vedere o diferență de cel puțin 40 pp, **în aproape jumătate dintre comune și o treime dintre orașe nu s-a pus deloc problema câștigătorului.** Considerând o diferență de minimum 20 pp, adică liniștitoare, **competiția aproape nu a existat în cazul a 69% dintre comune, respectiv 63% dintre orașe.**

Ponderea localităților în funcție de diferența dintre procentul voturilor obținute de primii doi candidați la funcția de primar



Calitatea estimărilor sondajelor: alegerile Europarlamentare 2024

Data: 2024-06-28

Cuvinte cheie: alegerile-europarlamentare-2024, performanta-sondajelor

În general calitatea sondajelor este evaluată în funcție de cât de apropiate sunt estimările obținute de rezultatele finale. Considerând principalii competitori, putem calcula media diferențelor absolute dintre estimări și rezultatele oficiale (eroarea medie). Putem ține cont de toți competitorii sau îi putem selecta pe aceia care îndeplinesc anumite criterii: toți cei care au obținut cel puțin 1/3/5% din voturi, respectiv primii 1/2/3 competitori în ordinea procentelor obținute.

O metodologie mai complexă de evaluare ar putea lua în calcul și sensul diferențelor observate, respectiv tipul și cantitatea informațiilor oferite cu privire la sondaj (transparența metodologică). Sensul diferențelor se referă la situația în care un institut supra-estimează, respectiv sub-estimează sistematic un anumit partid / candidat. Vorbim în acest caz de o eroare sistematică numită și distorsiune / părtinire (bias).

Transparența metodologică se referă de obicei la informațiile prezentate în caseta tehnică a sondajului: cum și când au fost culese datele, câte persoane au fost chestionate, care este reprezentativitatea teoretică: marja de eroare și nivelul de încredere. Pe lângă aceste întrebări, importante mai sunt cel puțin următoarele: cum a fost formulată întrebarea de vot și care au fost variantele de răspuns, dacă variantele de răspuns au fost citite sau

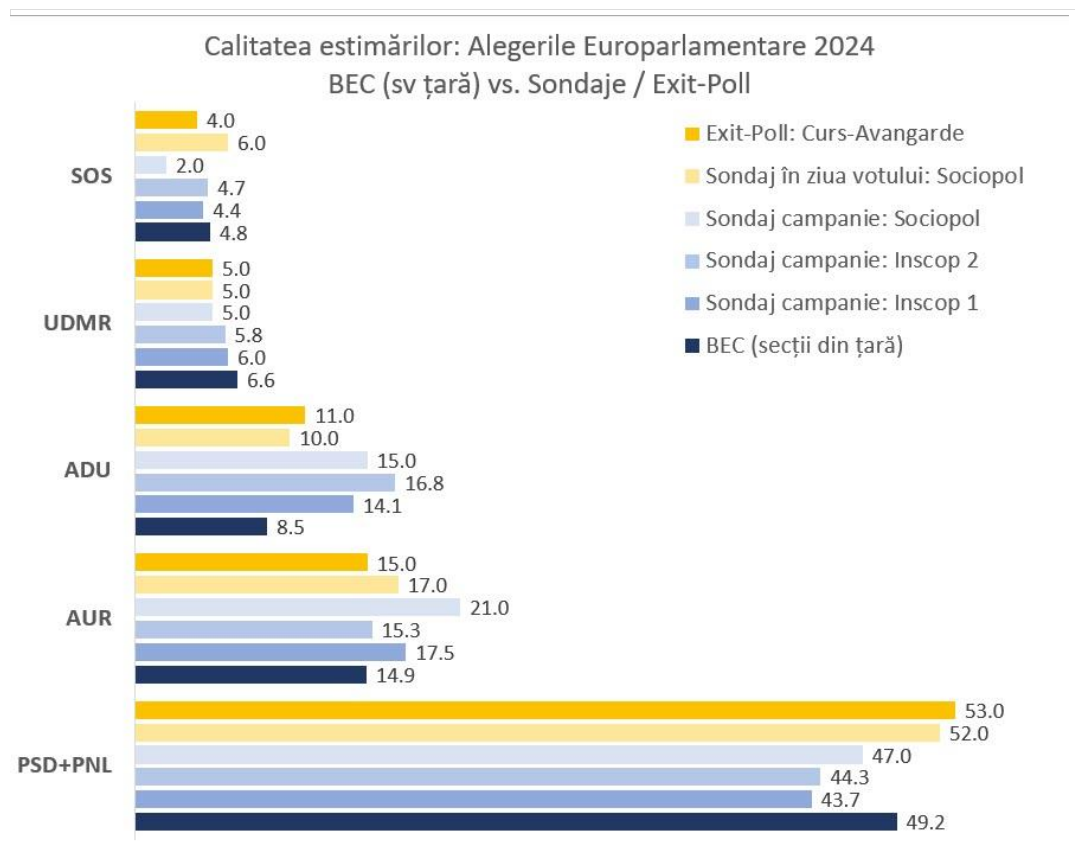
nu, în ce context a fost adresată întrebarea de vot, care a fost numărul răspunsurilor valide, care a fost cadrul de eşantionare şi cum a fost selectat efectiv eşantionul, care a fost rata de răspuns la anchetă, dacă datele au fost ponderate sau nu (cum anume), cine a finanţat sondajul etc.

Pentru această analiză avem în vedere doar criteriul eroare (medie). Vom calcula media erorilor pentru diferite situaţii: toţi candidaţii care au obţinut cel puţin 4% din voturi, primii trei, primii doi. Adăugăm la această listă eroarea asociată candidatului de pe prima poziţie, respectiv cea mai mare eroare înregistrată. Dat fiind faptul că toate sondajele au avut în vedere doar votanţii din ţară, ne-am raportat la rezultatele oficiale înregistrate în secţiile de vot din ţară.

Cu privire la alegerile Europarlamentare din 2024 am identificat în presă şi social-media doar patru sondaje. Două dintre aceste sondaje au fost realizate telefonic, în campania electorală, de către Inscop (au prezentat două estimări ale intenţiilor de vot: % din cei care au indicat un partid, respectiv % din cei care au declarat că sigur votează şi au indicat un partid) şi Sociopol. Celelalte două sondaje au fost realizate în ziua alegerilor: un sondaj la ieşirea de la urne (exit-poll Curs-Avangarde) şi un sondaj realizat, cel mai probabil, telefonic, în ziua votului (Sociopol). Cu excepţia sondajului Inscop, nu am găsit online (site-uri de ştiri, social-media, site-uri institute) informaţii despre metodologia celorlalte sondaje.

În general estimările intenţiilor de vot au fost consonante între institute, respectiv apropiate de rezultatul final. În cazul coaliţiei PSD+PNL sondajele din campanie au indicat un scor ceva mai mic, iar cele din ziua alegerilor un scor ceva mai mare, comparativ cu cel obţinut la urne. În cazul AUR sondajele au anticipat în general un scor mai mare decât cel obţinut. Estimarea produsă de sondajul Sociopol din campanie a supra-estimat cel mai mult scorul AUR. Cu privire la ADU, sondajele din campanie au indicat scoruri semnificativ mai mari decât cele obţinute, iar cele din ziua votului au fost apropiate de rezultatul final. O mare parte a acestei diferenţe poate fi explicată prin faptul că, spre finalul campaniei, o parte a electoratului ADU s-a reorientat spre candidaţii

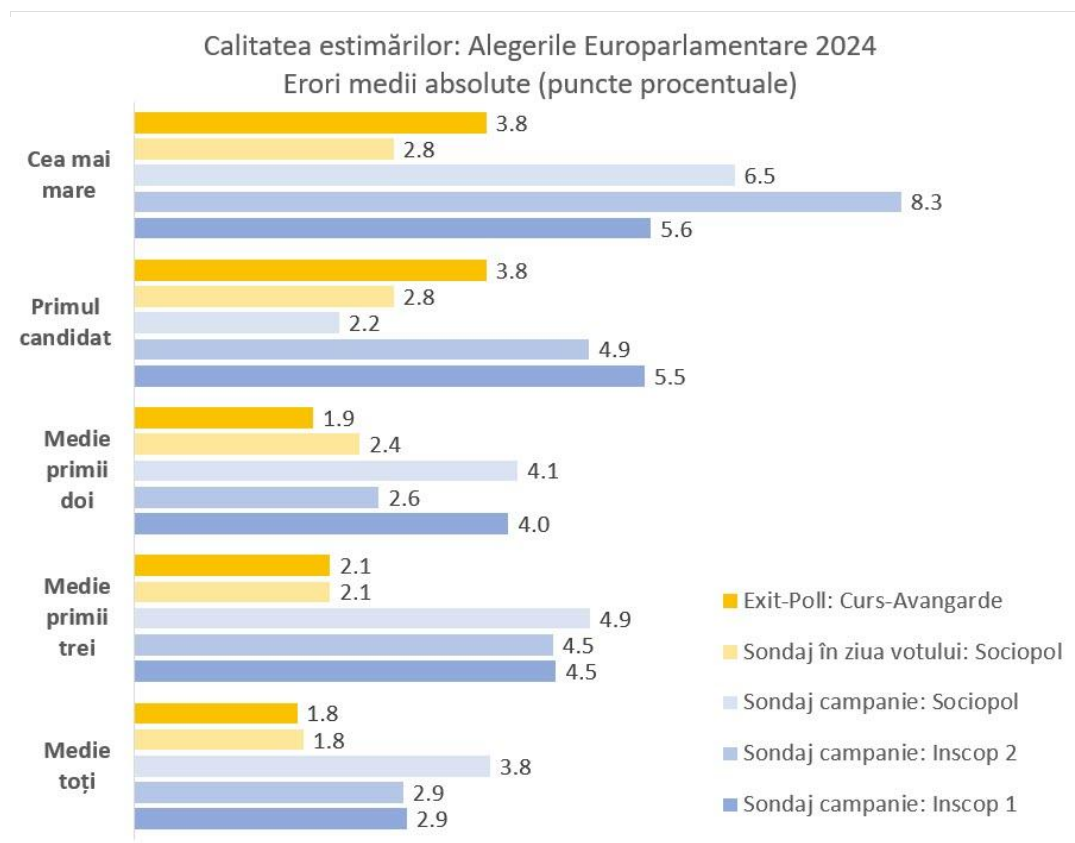
independenți. În cazul UDMR estimările au fost puțin sub rezultatul oficial, diferența fiind produsă în principal de o mobilizare mai mare decât de obicei a electoratului de etnie maghiară. La SOS estimările au fost din nou apropiate de rezultatul final (sondajele Sociopol au sub-estimat, respectiv supra-estimat puțin procentul obținut de SOS).



Dacă avem în vedere media diferențelor absolute pentru toți candidații care au obținut cel puțin 4%, sondajele au produs estimări foarte bune. Astfel, sondajele pre-electorale au avut o eroare medie de 3 puncte procentuale, iar cele din ziua votului de 1.8 puncte procentuale.

Dacă avem în vedere media erorilor în cazul primilor doi sau trei candidați, eroarea medie ajunge la 1.9-2.4 în cazul sondajelor din ziua votului, respectiv 4.0-4.9 în cazul sondajelor pre-electorale.

Dacă suntem și mai restrictivi și considerăm eroarea în cazul candidatului de pe prima poziție, respectiv cea mai mare eroare înregistrată, eroarea crește și mai mult. Din nou, eroarea este mai mare în cazul sondajelor pre-electorale. Scorul obținut de candidatului de pe prima poziție a fost estimat cel mai bine de sondajul Sociopol din campanie.



Erorile relativ mai mari observate în cazul sondajelor pre-electorale comparativ cu cele din ziua votului nu sunt deloc o surpriză (se întâmplă la fel și în alte țări) și pot fi explicate relativ simplu. O parte a alegătorilor se decid dacă și cu cine votează în ultima săptămână de campanie (10-15%) sau chiar în ziua votului (10-15%), iar cele două sondaje pre-electorale au fost realizate în prima parte a campaniei electorale (ultima zi de culegere a datelor a fost 22, respectiv 25 mai), deci cu două săptămâni înainte de vot. Sondajele pre-electorale au discutat cu puțin peste o mie de alegători, doar o parte a acestora fiind și votanți. Sondajele din

ziua votului au discutat cu mii de subiecți, iar în cazul sondajului la urne toți subiecții au fost și votanți. Foarte probabil rata de răspuns (la sondaj și la întrebarea de vot) a fost mai mare în cazul sondajelor din ziua votului. În cazul sondajelor din campanie măsurăm intenții de vot într-un orizont de timp de două-trei săptămâni), iar în ziua votului măsurăm comportamente de vot și / sau intenții de vot într-un orizont de timp de câteva ore.

Calitatea estimărilor sondajelor: alegerile pentru funcția de primar general

Data: 2024-07-13

Cuvinte cheie: performanta-sondajelor, alegeri-primar

Cele mai multe, mai vizibile și probabil mai controversate sondaje din această campanie au fost cele cu privire la funcția de primar general al Municipiului București. Subiectul acestei postări este analiza calității estimărilor acestor sondaje. Am inclus în analiză ultimul sondaj realizat în campanie de către fiecare organizație (dintre cele făcute publice), respectiv cele două sondaje realizate în ziua votului (Exit-Poll și CATI). Reamintim faptul că sondajele din campanie măsoară situația la acel moment și că, de obicei, o pondere relativ mare a alegătorilor se decid dacă și cu cine votează doar în ultima săptămână de campanie sau chiar în ziua votului. Totuși, dacă un sondaj din campanie prezintă estimări foarte diferite de celelalte sondaje și simultan foarte diferite de rezultatul oficial, sunt șanse mari ca acel sondaj să fi fost realizat incorect, respectiv să fi fost o încercare de manipulare.

Rezultatele oficiale (BEC) și estimările produse de sondajele incluse în analiză sunt prezentate în tabelul de mai jos. Estimările au fost preluate din media online, respectiv site-urile institutelor (în cazul sondajului CATI realizat de Sociopol în ziua alegerilor, deoarece nu am găsit estimarea pentru candidatul Enache, am trecut diferența până la 100%, adică 5%). În general, atunci când au fost făcute publice, casetele tehnice ale sondajelor s-au limitat strict la informațiile impuse prin lege. Cu privire la sondajul CATI

din ziua alegerilor realizat de Sociopol nu am găsit online caseta tehnică. Pe site-ul CURS a apărut ulterior un [comentariu](#) cu privire la Exit-Poll (acesta include și o descriere a metodologiei și resurselor logistice utilizate). Acronimele care însoțesc numele institutelor se referă la modalitatea prin care au fost colectate datele: CATI = sondaj telefonic, F2F = față în față, INTERNET = „recrutare (aleatoare) digitală” în cazul Atlasintel și „online” în cazul Epsilon.

Primar General	Dan	Firea	Piedone	Burduja	Enache
BEC	47.9	22.2	15.1	7.8	3.0
Ultimul sondaj din campanie:					
Jarding (CATI)	27.2	19.3	17.7	27.8	5.4
Avangarde (F2F)	28.6	26.5	15.2	13.8	7.4
CURS (F2F)	30.0	29.0	16.0	13.0	7.0
Epsilon (INTERNET)	31.0	32.0	11.0	9.0	4.0
Sociopol (CATI)	31.0	26.0	15.0	16.0	7.0
ARA (CATI)	35.0	29.0	20.0	12.0	3.0
IPSOS (F2F)	43.0	19.1	27.6	5.7	1.9
IRSOP (CATI)	43.0	18.0	23.0	4.0	0.0
Atlasintel (INTERNET)	45.2	21.7	15.4	7.8	5.7
Sondaj în ziua votului:					
Exit-Poll: Curs-Avangarde	44.0	25.0	12.0	10.0	5.0
CATI: Sociopol	47.0	24.0	14.0	10.0	5.0

Similar cu [analiza anterioară](#), și în acest caz am luat în calcul doar criteriul diferență absolută (medie) relativ la rezultatele oficiale. Am calculat media diferențelor absolute pentru diferite situații: toți candidații care au obținut cel puțin 3% din voturi, primii trei, primii doi. Am adăugat la această listă diferența asociată candidatului de pe prima poziție, respectiv cea mai mare diferență înregistrată. Rezultatele obținute sunt prezentate în tabelul de mai jos, ordinea de prezentare fiind descendentă în funcție de mărimea diferențelor.

Diferență absolută medie	Toți	Primii trei	Primii doi	Primul candidat	Cea mai mare
Ultimul sondaj din campanie:					
Jarding (CATI)	9.7	8.7	11.8	20.7	20.7
Avangarde (F2F)	6.8	7.9	11.8	19.3	19.3
CURS (F2F)	7.0	8.5	12.4	17.9	17.9
Epsilon (INTERNET)	6.6	10.3	13.4	16.9	16.9
Sociopol (CATI)	6.6	7.0	10.4	16.9	16.9
ARA (CATI)	5.8	8.2	9.9	12.9	12.9
IPSOS (F2F)	4.7	6.8	4.0	4.9	12.5
IRSOP (CATI)	4.8	5.7	4.6	4.9	7.9
Atlasintel (INTERNET)	1.2	1.2	1.6	2.7	2.7
Sondaj în ziua votului:					
Exit-Poll: Curs-Avangarde	2.8	3.3	3.4	3.9	3.9
CATI: Sociopol	1.6	1.3	1.4	0.9	2.2

Așa cum era de așteptat, sondajele realizate în ziua votului au produs cele mai bune estimări. Dintre cele două estimări, estimarea CATI (Sociopol) a fost mai apropiată de rezultatul oficial comparativ cu estimarea de la Exit-Poll (Curs-Avangarde), indiferent ce criteriu, dintre cele incluse aici, considerăm. La prima vedere, dată fiind diferența dintre volumele celor două eșantioane (15000 la Exit-Poll și probabil 1-2000 la CATI), o calitate mai bună a estimărilor de către sondajul CATI poate fi considerată o surpriză. În realitate, pentru motive ce țin de diferențele tehnice dintre cele două tipuri de sondaje, surpriza nu este atât de mare. În cazul unui sondaj la urne (Exit-Poll) nu se pune problema identificării votanților, numărul de chestionare aplicate este semnificativ mai mare, dar, pentru că votanții sunt selectați dintr-un eșantion de secții de vot, eșantionul efectiv e în realitate semnificativ mai mic decât numărul de persoane care au participat la sondaj. În cazul unui sondaj CATI problema identificării votanților rămâne, numărul de chestionare este relativ mai mic, dar subiecții sunt selectați direct din (aproximativ) toată populația vizată de sondaj. De exemplu, pentru acest Exit-Poll, un număr de 15000 votanți selectați din 140 secții de vot este echivalent cu aproximativ 1500 votanți selectați simplu aleator (funcție de cât de de similari între ei sunt

votanții dintr-o secție de vot oarecare). Prin urmare, un eșantion simplu aleator de 1500 votanți, așa cum probabil a fost eșantionul CATI din ziua alegerilor, poate fi la fel de bun ca un eșantion de 15000 colectat la un Exit-Poll. Pentru o explicație mai elaborată a acestui „paradox” se poate consulta și postarea [SRS vs. Cluster](#).

În cazul sondajele din campanie, diferențele față de rezultatul final au fost semnificativ mai mari, cu o excepție notabilă: sondajul Atlasintel realizat online prin tehnica „recrutare digitală aleatoare.” Desigur, o astfel de tehnică poate produce rezultate bune doar dacă populația este una digitalizată, cum de altfel este și cazul populației din Municipiului București. Aplicând aceeași tehnică, de această dată la nivel național (alegerile euro-parlamentare din 2024, România), același institut a produs estimări foarte diferite de cele oficiale: USR a fost supra-estimat cu 16 puncte procentuale (pp), iar coaliția PSD-PNL a fost sub-estimată cu aproape 18 pp.

În ordinea descrescătoare a calității estimărilor sondajelor din campanie, pe poziția secundă se află sondajele realizate de IPSOS și IRSOP. Pe locul trei se situează, cu unele diferențe între ele, sondajele realizate de celelalte institute. Toate aceste sondaje au în comun faptul că au sub-estimat foarte mult (13-21 pp) votul pentru Nicușor Dan. Desigur, o parte din diferența observată este explicată de faptul că datele au fost colectate cu una până la patru săptămâni înainte de ziua alegerilor, la acel moment intențiile de vot fiind foarte probabil diferite de rezultatele oficiale.

Cele mai mari diferențe apar în cazul sondajului realizat de compania Jarding: Nicușor Dan a fost sub-estimat cu aproape 21 pp, iar Sebastian Burduja a fost supra-estimat cu 20 pp. Aceste diferențe sunt cu atât mai neobișnuite cu cât datele au fost colectate într-un interval de 3-7 zile înainte de ziua alegerilor, iar restul sondajelor indicau o estimare foarte diferită pentru candidatul Burduja.

Sondajele incluse în această evaluare diferă în funcție de multe caracteristici, modalitatea concretă (tehnica) de colectare a datelor de sondaj (F2F, CATI, INTERNET) fiind doar una dintre acestea. În plus, numărul de sondaje din fiecare categorie este foarte mic. Mai mult, nu știm dacă estimările publicate sunt și cele

rezultate din teren, iar dacă diferă, nu știm cum s-a ajuns la estimările publicate. Cu aceste limite, putem totuși compara, orientativ, diferențele de calitate a estimărilor în funcție de tehnica de sondaj. Se observă că, în cazul unei populații digitalizate, media estimărilor produse de cele două sondaje realizate online (INTERNET) este ceva mai bună. Foarte important, aceasta se întâmplă doar pentru că unul dintre sondajele online (Atlasintel) a fost foarte aproape de rezultatele finale (cel realizat de Epsilon a fost la mare distanță). Calitatea medie a estimărilor produse de sondajele F2F este aproximativ aceeași cu calitatea medie a estimărilor produse de sondajele CATI.

Diferență absolută medie	Toți	Primii trei	Primii doi	Primul candidat	Cea mai mare
F2F	6.2	7.8	9.4	14.1	16.6
CATI	6.7	7.4	9.2	13.9	14.6
INTERNET	3.9	5.7	7.5	9.8	9.8

AUR + Sociopol = Love?

Data: 2024-10-26

Cuvinte cheie: alegeri-locale, performanta-sondajelor

Continui seria de postări pe tema evaluării calității sondajelor de la alegerile locale 2024 cu o scurtă analiză a estimărilor intențiilor de vot în cadrul unor competiții electorale locale mai puțin vizibile la nivel național (funcția de Primar, Președinte al Consiliului Județean, Consilier Local din câteva localități), așa cum apar acestea în sondajele Sociopol. Estimările au fost preluate de pe paginile de Facebook ale lui Mirel Palada și Ovidiu Voicu. Am redus analiza doar la Sociopol pentru un motiv foarte simplu: este singurul institut care a publicat estimări pentru o serie de competiții locale. Estimările din sondaje, rezultatele oficiale și diferența dintre acestea sunt prezentate în tabelele de la finalul postării.

Deși nu am avut date despre volumul eșantioanelor (sau metodologia sondajelor), putem presupune că fiecare eșantion a fost de aproximativ 1000 subiecți (poate ceva mai puțin), prin urmare eroarea de eșantionare a fost de aproximativ 4 puncte procentuale (am ținut cont și de faptul că nu toți cei 1000 subiecți au și declarat intenția de vot). Observăm că media diferențelor absolute dintre estimări și datele oficiale tinde să fie aproape dublă față de eroarea de eșantionare proiectată. Acest lucru nu este surprinzător și poate fi observat în foarte multe alte situații similare: în general, estimările de vot obținute în sondaje sunt mai apropiate de rezultatele oficiale atunci când populația de interes este cea de la nivel național comparativ cu cea de la nivel local.

Ceea ce surprinde însă este distorsiunea estimărilor: **sistematic, candidații AUR sunt supra-estimați. Supra-estimarea medie este de aproape 15 puncte procentuale.** În niciunul dintre sondaje nu a apărut o situație de sub-estimare a candidaților AUR. Dat fiind faptul că alegerile analizate sunt pentru funcții diferite, respectiv la nivelul unor localități diferite, situația observată poate fi explicată doar prin acțiunea unui factor care a acționat sistematic. Acest factor poate fi de natură metodologică sau nu. Prin metodologie avem în vedere colectarea, ponderarea și modelarea datelor (asumpțiile din spatele modelelor utilizate, în cazul în care au fost utilizate astfel de modele). Pentru context, la alegerile pentru Parlamentul European 2024, Sociopol a supra-estimat puțin intenția de vot pentru AUR în sondajele naționale din campanie, respectiv aproape deloc în sondajul telefonic din ziua votului.

Diferență medie ...	AUR	RESTUL
pozitivă	14.8	3.5
negativă	0.0	-8.1
absolută	14.8	6.5

Primar: Sector 1	Partid	Sociopol	BEC	Diferență
Armand (ADU)	ADU	32	35.9	-3.9
Tuță (PSD-PNL)	PSD-PNL	28	36.6	-8.6
Tudorache	INDEP	15	7.6	7.4
Bruyseels (AUR)	AUR	14	7.1	6.9
Negoită (PU-SL)	PU-SL	8	7.0	1.0

Primar: Iași	Partid	Sociopol	BEC	Diferență
Chirică (PNL)	PNL	29	32.2	-3.2
Balanîșcu (PSD)	PSD	22	13.0	9.0
Bodea (ADU)	ADU	18	26.1	-8.1
Ciuhodaru (AUR)	AUR	14	8.9	5.1
Chichirău (R-Eco)	R-Eco	11	14.9	-3.9

Primar: Cluj-Napoca	Partid	Sociopol	BEC	Diferență
Boc (PNL)	PNL	33	42.5	-9.5
Cozma (AUR)	AUR	28	3.6	24.5
Sărmaș	INDEP	14	29.4	-15.4
Cristea (PSD)	PSD	7	4.6	2.4
Băltărețu (ADU)	ADU	6	5.9	0.1
Olah (UDMR)	UDMR	6	9.4	-3.4

Primar: Suceava	Partid	Sociopol	BEC	Diferență
Rîmbu (PSD)	PSD	27	42.5	-15.5
Axinte (AUR)	AUR	25	9.3	15.7
Harșovschi (PNL)	PNL	20	27.7	-7.7
Andronache (ADU)	ADU	12	15.4	-3.4
Neșculescu (SOS)	SOS	7	2.4	4.6
Harasim (PU-SL)	PU-SL	6	0.8	5.2

Primar: Botoșani	Partid	Sociopol	BEC	Diferență
Andrei (PSD)	PSD	30	58.0	-28.0
Flutur (PNL)	PNL	28	25.1	2.9
Silegeanu (AUR)	AUR	22	6.4	15.6
Jurgiu (SOS)	SOS	8	2.3	5.7
Curelariu (ADU)	ADU	7	4.4	2.6

Primar: Focșani	Partid	Sociopol	BEC	Diferență
Misăilă (PSD)	PSD	30	51.1	-21.1
Anghene (AUR)	AUR	26	7.9	18.1
Resmeriță (PNL)	PNL	23	21.4	1.6
Dimitriu (ADU)	ADU	9	11.4	-2.4
Pleșa (SOS)	SOS	6	2.8	3.2

PCJ: Ilfov	Partid	Sociopol	BEC	Diferență
Thuma (PSD-PNL)	PSD-PNL	41	59.3	-18.3
Mihalcea (AUR)	AUR	34	13.0	21.0
Dancă (ADU)	ADU	18	19.3	-1.3
Bobei (SOS)	SOS	4	5.7	-1.7

PCJ: Vrancea	Partid	Sociopol	BEC	Diferență
Halici (PSD)	PSD	34	37.7	-3.7
Badiu (AUR)	AUR	31	9.6	21.4
Ciobotaru (PNL)	PNL	26	36.8	-10.8
Ștefan (ADU)	ADU	4	10.4	-6.4

PCJ: Botoșani	Partid	Sociopol	BEC	Diferență
Trufin (PSD)	PSD	38	41.7	-3.7
Iftime (PNL)	PNL	28	41.9	-13.9
Cilibiu (AUR)	AUR	20	8.6	11.4
Amarandei (SOS)	SOS	8	2.8	5.2
Lăcătușu (ADU)	ADU	4	3.2	0.8

CL: Cluj-Napoca	Partid	Sociopol	BEC	Diferență
	PNL	28	32.4	-4.4
	ADU	19	18.4	0.6
	AUR	17	8.8	8.3
	PSD	12	11.1	0.9
	UDMR	11	14.2	-3.2
	REPER	7	8.1	-1.1

Turul 2 al alegerilor prezidențiale din Moldova: o modelare a intențiilor de vot

Data: 2024-10-26

Cuvinte cheie: alegerile-prezidentiale-2024-moldova,
modelarea-intențiilor-de-vot

Calitatea estimărilor intențiilor de vot obținute de sondajele de opinie variază destul de mult de la o țară la alta. Printre factorii responsabili de aceste diferențe se numără adesea gradul de democrație / libertate, nivelul de sărăcie, tipul de cultură, confluența intereselor socio-politice internaționale, respectiv numărul alegătorilor. Într-o țară predominant săracă, cu un număr mic de alegători, situată într-o zonă geografică sensibilă, așa cum este Moldova, nu e de mirare că sondajele produc uneori o imagine deformată a realității prezente sau viitoare. În astfel de contexte, pentru a anticipa rezultatele electorale, e obligatoriu să folosim, alături de date de sondaj, și alte informații (adesea la nivel agregat), precum rezultatele recente ale alegerilor.

Propun aici un astfel de exercițiu bazat strict pe datele oficiale de la turul 1 și o intuiție educată parțial (nu sunt un cunoscător al spațiului politic din Moldova, al afinităților dintre partide, candidați, relațiilor dintre zona economică și politică, respectiv dintre instituțiile de forță și cele politice). Cu siguranță acest exercițiu ar avea sorti mai mari de reușită dacă ar beneficia de sprijinul unor experți locali, respectiv dacă ar avea acces la date de sondaj care să măsoare opțiunea de vot de la turul 1 și intenția de vot pentru turul 2. Desigur, o astfel de modelare prezintă o serie

de alte limite. Una dintre acestea e legată de o posibilă schimbare (de la turul 1 la turul 2) a direcției intensității presiunilor exercitate de actorii majori asupra alegătorilor. Cu cât direcția combinată a presiunilor va fi mai diferită comparativ cu turul 1, cu atât impactul asupra rezultatului final va fi mai mare. Alte două limite privesc influența dezbaterilor electorale (ele urmează să aibă loc) și alianțele politice negociate probabil în această perioadă undeva departe de lumina reflectoarelor.

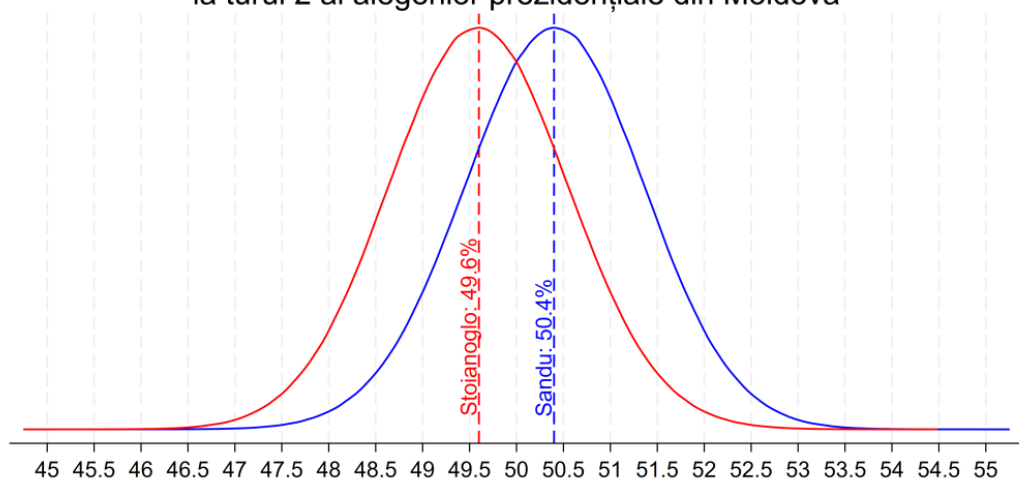
Pe lângă datele oficiale, modelarea se bazează pe două tipuri de cantități:

- transferul de voturi dinspre candidații eliminați spre cei rămași în cursă, respectiv numărul alegătorilor care au votat la turul 1 dar nu vor vota la turul 2;
- numărul de votanți noi (cei care nu au votat la turul 1 dar vor vota la turul 2) atrași de fiecare dintre cei doi finaliști.

Cum foarte probabil situația va fi diferită în funcție de context – țară vs. diaspora – am folosit valori diferite pentru cele două populații. Desigur, ar fi fost utilă o separare mai fină în funcție de raion, respectiv de țara în care se află secțiile de vot din străinătate.

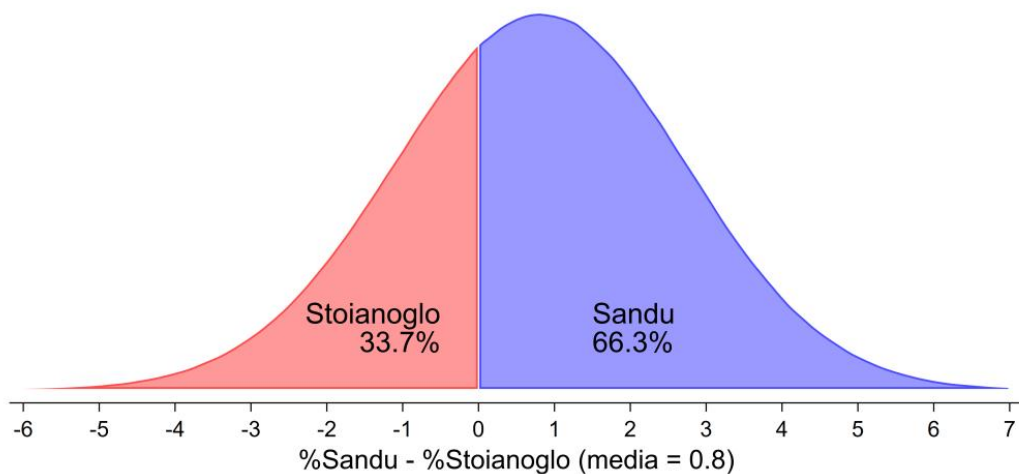
Rezultatele obținute în urma simulărilor indică o competiție foarte strânsă pentru turul 2. Conform modelului, distanța dintre cei doi candidați va fi probabil de aproximativ două puncte procentuale. În setul de scenarii care a stat la baza acestei simulări, diferența medie dintre candidați a fost de 0.8 puncte procentuale în favoarea Maiei Sandu. Probabilitatea ca Maia Sandu să câștige alegerile este însă de doar 66%. Simplu spus, considerând scenariul de bază definit de valorile asociate celor două tipuri de parametri menționați anterior, modelul estimează că Maia Sandu va câștiga 66 din 100 alegeri.

Estimarea intențiilor de vot la turul 2 al alegerilor prezidențiale din Moldova



Notă: Pentru estimare am folosit rezultatele de la T1 și o definiție subiectivă a situației așteptate la T2: redistribuirea voturilor candidaților eliminați și numărul noilor votanți.

Estimarea probabilității de a câștiga turul 2 al alegerilor prezidențiale din Moldova



Notă: Pentru estimare am folosit rezultatele de la T1 și o definiție subiectivă a situației așteptate la T2: redistribuirea voturilor candidaților eliminați și numărul noilor votanți.

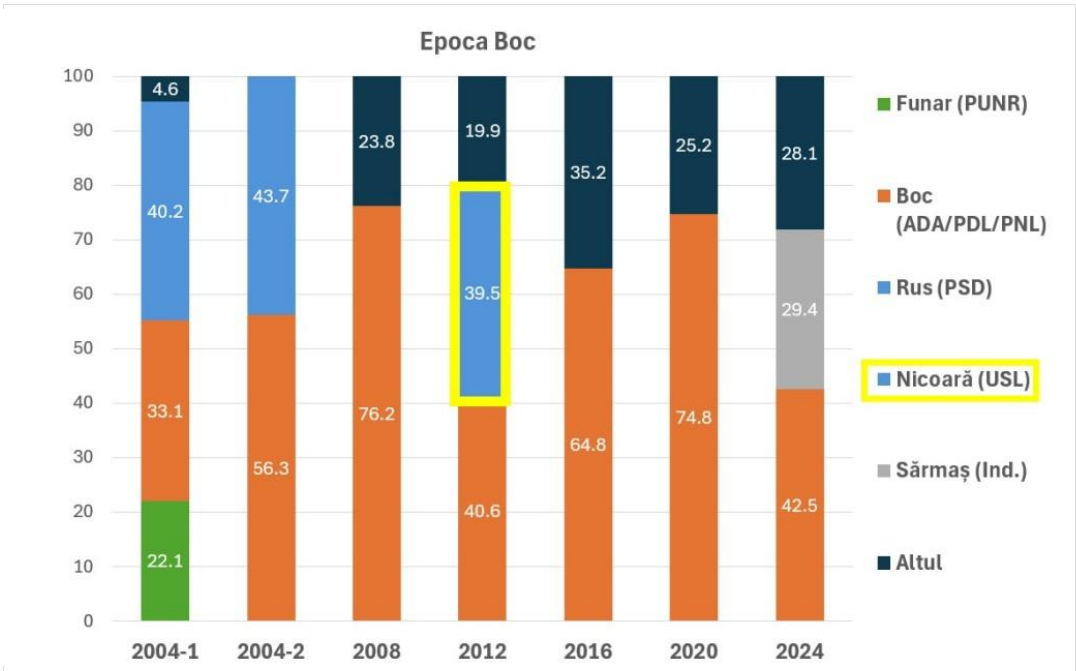
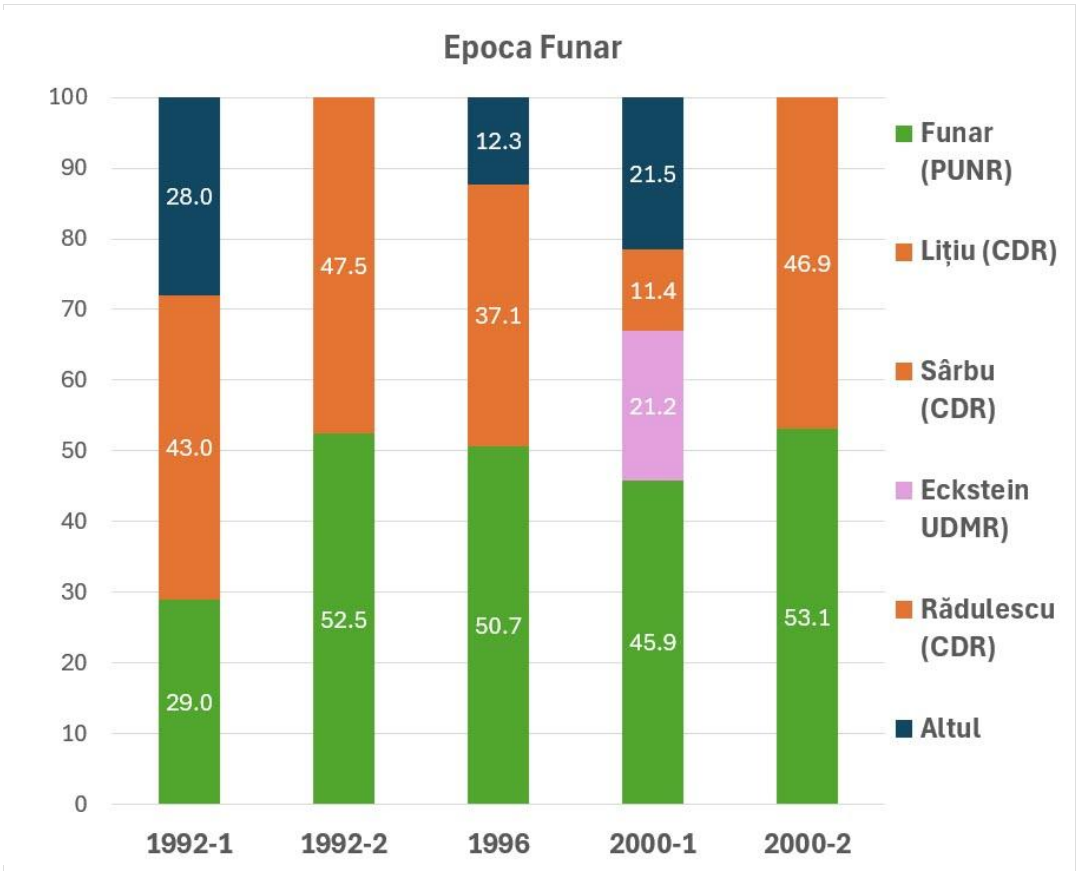
Un oraș, doi primari

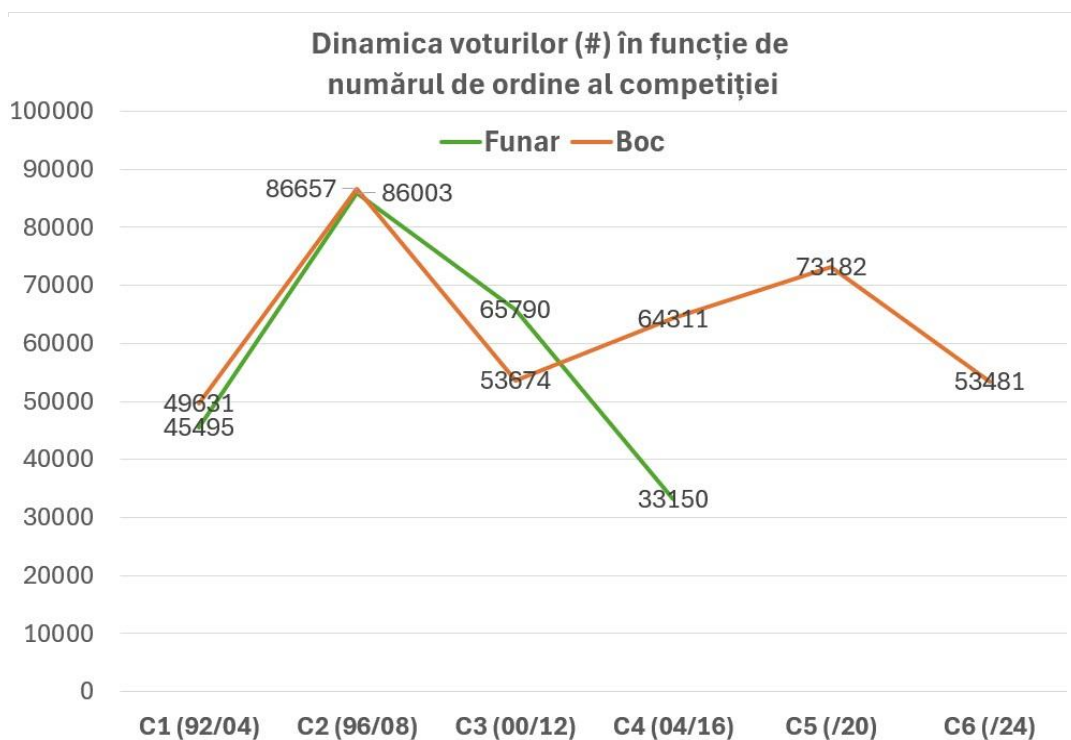
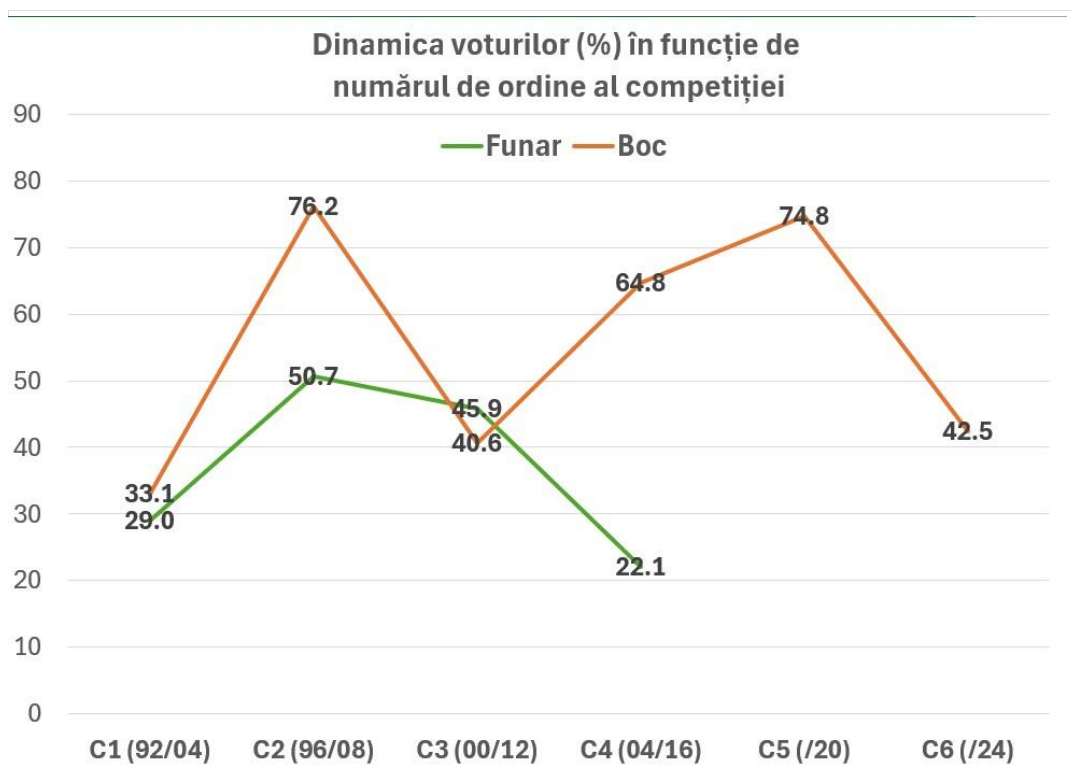
Data: 2024-10-26

Cuvinte cheie: alegeri-locale, alegerile-locale-2024, cluj-napoca, alegeri-primar

Cluj-Napoca este una dintre puținele localități mari care, după 1990, au avut doar doi primari. În această scurtă postare prezint câteva grafice care compară, din punct de vedere electoral, cele două epoci: Funar și Boc. Pe baza acestora observăm că:

- Niciunul dintre cei doi primari nu a pornit ca favorit în prima competiție electorală. Însă, la turul doi, fiecare dintre ei a reușit să atragă mai multe dintre voturile candidaților eliminați, respectiv să aducă mai mulți votanți noi la urne.
- Chiar dacă Funar a câștigat următoarele două mandate, competiția a fost strânsă de fiecare dată. Boc a avut parte de competiții strânse doar atunci când a existat un contra-candidat real (Nicoară, Sărmaș).
- Dinamica opțiunilor pentru Funar ia forma literei V inversate. În cazul lui Boc, ia forma literei M. Oare la următoarea competiție vom asista la o revenire sau la o continuare a scăderii (sau niciuna dintre acestea pentru că actualul primar se va retrage strategic pentru a nu periclita șansele PNL)? Totul depinde în principal de doi factori: (in)succesul proiectelor majore derulate pe parcursul noului mandat, respectiv de „greutatea electorală” a viitorului contra-candidat.





Turul 2 al alegerilor prezidențiale din Moldova: impactul posibil al fenomenului de cumpărare a voturilor

Data: 2024-10-28

Cuvinte cheie: alegerile-prezidentiale-2024-moldova,
modelarea-intentiilor-de-vot

Se pare că numărul cazurilor de cumpărare a votului la turul 1 este foarte mare, undeva în jur de 150 mii, conform datelor oferite de [instituțiile statului](#). E posibil chiar ca acest număr să mai crească în următoarele zile. La prima variantă a modelării intențiilor de vot pentru turul 2 nu am luat deloc în calcul acest lucru (estimarea de 150 mii voturi cumpărate la turul 1 a apărut ieri). Un număr atât de mare al cazurilor de acest tip, investigațiile desfășurate în săptămâna care a trecut precum și modul în care au fost acestea au fost prezentate în media pot naște reacții diferite la nivelul populației.

Efectul poate fi unul important mai ales la nivelul alegătorilor implicați în această practică ilegală, dar și a celor care intenționau să voteze doar la turul 2. Pe de o parte, e relativ mai probabil ca o parte a celor care au primit bani pentru a vota la turul 1 cu NU, respectiv pentru a vota cu un alt candidat decât Maia Sandu și care ar fi votat la turul 2 cu Stoianoglo, să nu se mai prezinte la vot (situația A). Similar, o parte din cei care intenționau să voteze doar la turul 2 și care o preferau pe Maia Sandu relativ mai probabil nu vor mai vota la turul 2 (situația B). E dificil de estimat câți alegători

se află în fiecare dintre aceste două situații. Putem însă ghici, definind fiecare cantitate printr-un interval de variație relativ mai larg (pentru că incertitudinea asociată este mult mai mare în acest caz).

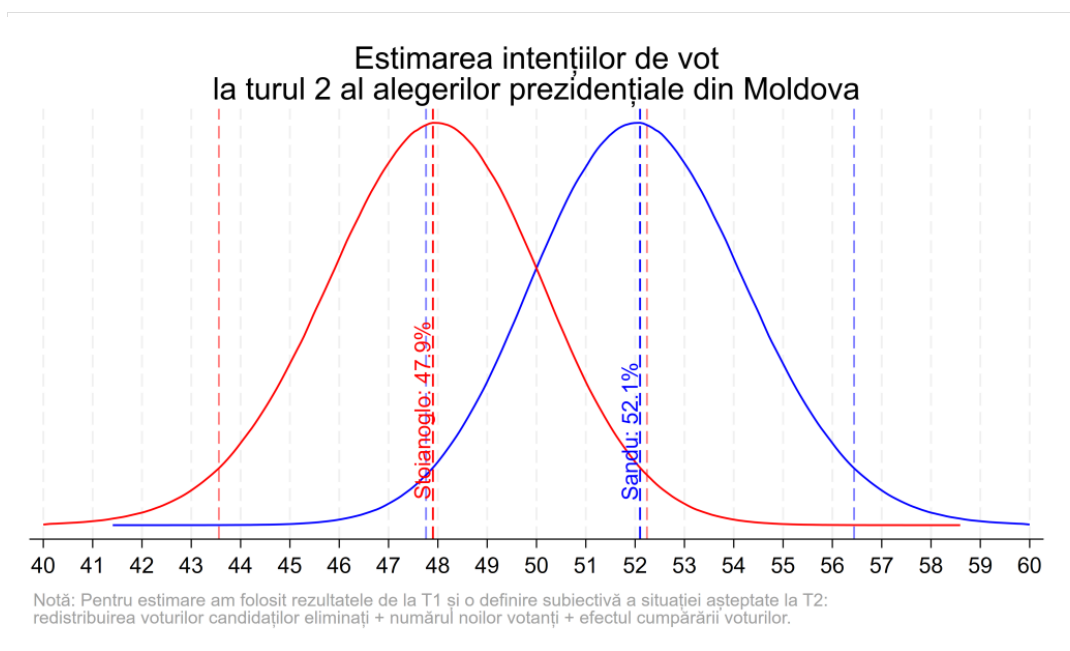
Situația A: Numărul cazurilor de cumpărare a voturilor e probabil mai mare decât cel raportat până la acest moment, să zicem că undeva spre 200 mii. Să presupunem că numărul celor care nu vor mai vota la turul 2 urmează o distribuție normală care acoperă 95% din situațiile posibile luând valori în intervalul $-215; 75 - 215; 75$ mii (deci o distribuție normală cu media -75 mii și o abatere standard 75 mii). După cum se observă, am inclus în acest scenariu și posibilitatea, deloc neglijabilă, ca fenomenul de cumpărare a voturilor să fie chiar ceva mai răspândit la turul 2 (intervalul de variație include și valori pozitive). Această distribuție descrie numărul de voturi de care trebuie adunate la „zestrea” lui Stoianoglo.

Situația B: E foarte dificil de ghicit cam câți alegători care nu au votat la turul 1 și care intenționau inițial să voteze la turul 2 cu Maia Sandu vor avea o reacție de retragere și nu vor mai vota. Și în acest caz incertitudinea este foarte mare, mai ales că nu există o valoare de care să ne ancorăm estimarea. În plus, există o oarecare posibilitate să apară un fenomen contrar, de supra-mobilizare a votanților pro Sandu. Să presupunem că acest număr ia, cel mai probabil, valori în intervalul $-75; 37,5 - 75; 37,5$ mii. Prin urmare, distribuția care descrie această situație este una normală cu media $-37,5$ mii și abaterea standard 37,5 mii. Această distribuție descrie numărul de voturi de care trebuie adunate la „zestrea” Maiei Sandu.

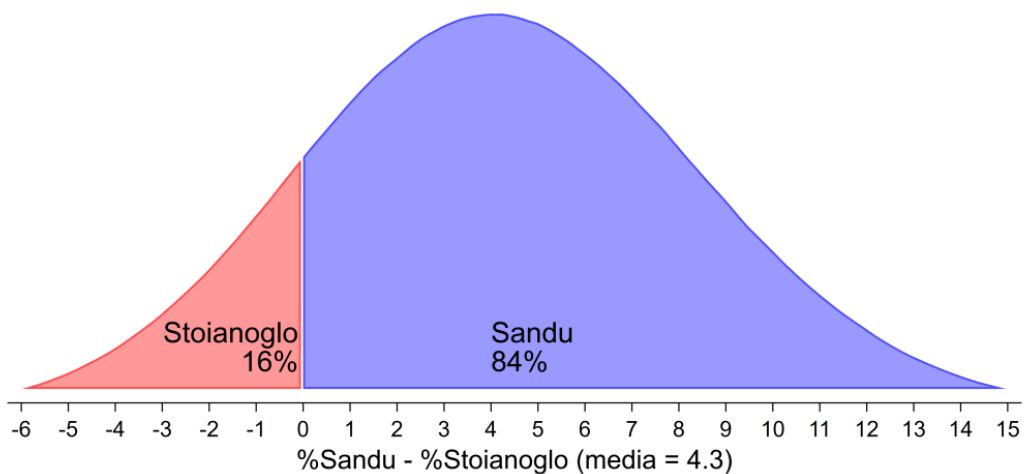
Pornind de la modelul inițial și luând în calcul noile evenimente, exprimate cantitativ prin distribuțiile statistice descrise mai sus, putem reface predicția descrisă în postarea anterioară. Varianta actualizată a modelării arată că probabilitatea ca Maia Sandu să câștige turul 2 este acum semnificativ mai mare: 96% față de 66% anterior. Conform acestui scenariu, așteptarea e ca Maia Sandu să obțină aproximativ 50–54% dintre voturile valabil exprimate.

Dat fiind faptul că nu avem date de sondaj sau de alt tip care să ne spună care vor fi reacțiile cele mai probabile ale alegătorilor, e

posibil ca scenariile pe care le-am imaginat aici să nu descrie suficient de bine viitoarea realitate, prin urmare predicțiile făcute pe baza acestor scenarii se vor dovedi greșite. Suplimentar, rezultatul final depinde și de posibilele alte evenimente semnificative care ar putea să apară în această săptămână. Destul de probabil, dezbaterile electorale nu vor fi unul dintre acești factori (literatura de specialitate arată că influența dezbaterilor, în termeni absoluți, e mai degrabă redusă).



Estimarea probabilității de a câștiga turul 2 al alegerilor prezidențiale din Moldova



Notă: Pentru estimare am folosit rezultatele de la T1 și o definiție subiectivă a situației așteptate la T2: redistribuirea voturilor candidaților eliminați + numărul noilor votanți + efectul cumpărării voturilor.

Turul 2 al alegerilor prezidențiale din Moldova: impactul posibil al ultimelor evenimente

Data: 2024-10-31

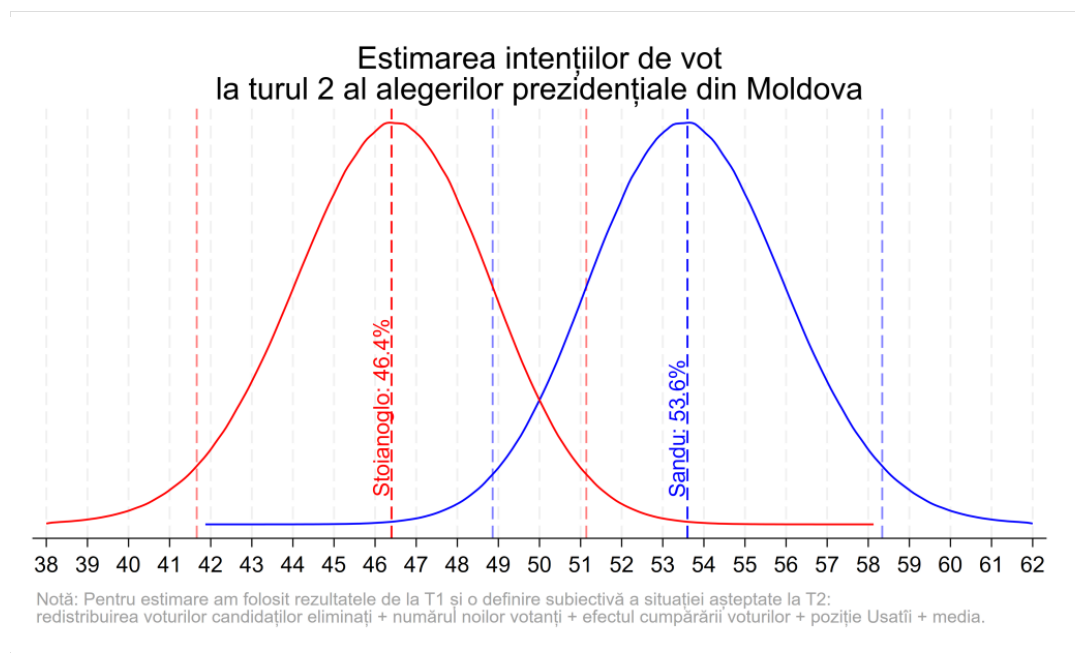
Cuvinte cheie: alegerile-prezidentiale-2024-moldova, modelarea-intentiilor-de-vot

Au trecut trei zile de la ultima modelare și mai sunt patru până la alegerile de duminică. În mod normal nu se mai pot întâmpla prea multe lucruri care să afecteze semnificativ distribuția voturilor. Totuși, între timp au avut loc două evenimente care ar putea schimba puțin distribuția voturilor de la turul 2.

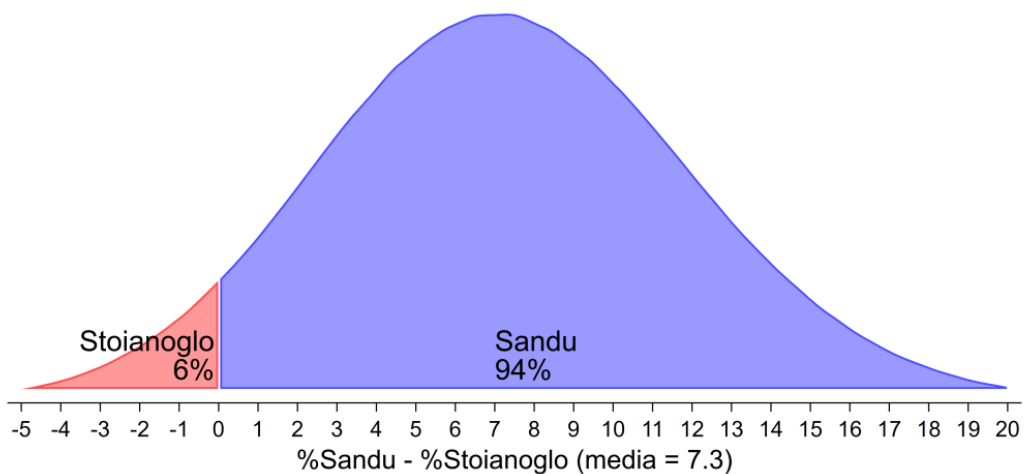
Usatîi a avut o [intervenție publică](#) în care i-a criticat în egală măsură pe cei doi finaliști și partidele lor, respectiv a susținut explicit că nu îl susține pe niciunul dintre aceștia. Ceva mai important pentru modelarea propusă, Usatîi a denunțat presiunile la care a fost supus, la început cu scopul de a-l determina să nu mai candideze, ulterior pentru a susține public că îl sprijină pe Stoianoglo. Estimez că efectul net asupra votanților ar putea fi ușor în favoarea Maiei Sandu, parțial și ca urmare a creșterii ponderii celor care nu vor vota la turul 2. În prima modelare am redistribuit voturile lui Usatîi în favoarea lui Stoianoglo (65% Stoianoglo, 20% Sandu, 15% non-votanți). În urma acestor declarații cred că distribuția va fi relativ mai echilibrată (50%, 25%, 25%). Reamintesc faptul că toate aceste valori sunt medii ale unor distribuții normale și că nu au la bază informații din sondaje (acestea lipsesc sau nu sunt publice).

Tot în ultimele zile, au fost prezentate în media o serie de cazuri de cumpărare a voturilor, au fost descrise măsurile luate de instituțiile responsabile (CNA), au fost menționate pedepsele stipulate de lege, s-a discutat despre faptul că instituțiile statului vor renunța la acuzații și amenzi în cazul în care cei implicați recunosc fapta și colaborează cu instituțiile statului, cu condiția să nu mai încalce legea și la turul 2. Estimez că toate acestea vor duce probabil la o scădere și mai mare a cazurilor de cumpărare a voturilor, respectiv la o prezență la vot mai redusă decât cea anticipată anterior.

În concluzie, comparativ cu modelarea anterioară, distanța dintre cei doi candidați este probabil în prezent ceva mai mare decât în urmă cu trei zile: 7 puncte procentuale față de 4. Evident, probabilitatea estimată ca Maia Sandu să câștige alegerile a crescut și ea, ajungând la 94% (anterior a fost 84%).



Estimarea probabilității de a câștiga turul 2 al alegerilor prezidențiale din Moldova



Notă: Pentru estimare am folosit rezultatele de la T1 și o definiție subiectivă a situației așteptate la T2:
redistribuirea voturilor candidaților eliminați + numărul noilor votanți + efectul cumpărării voturilor + poziție Usatii + media.

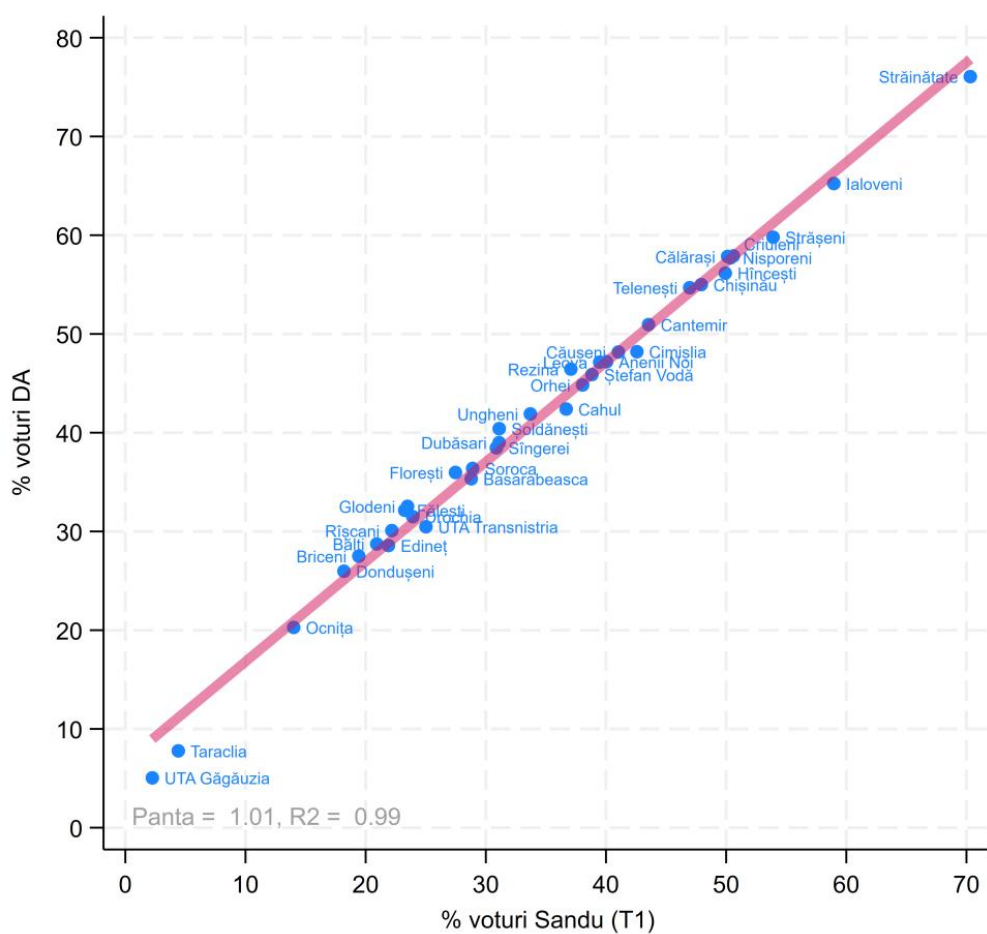
Alegeri 2024 Moldova: o analiză la nivel de raion a rezultatelor oficiale

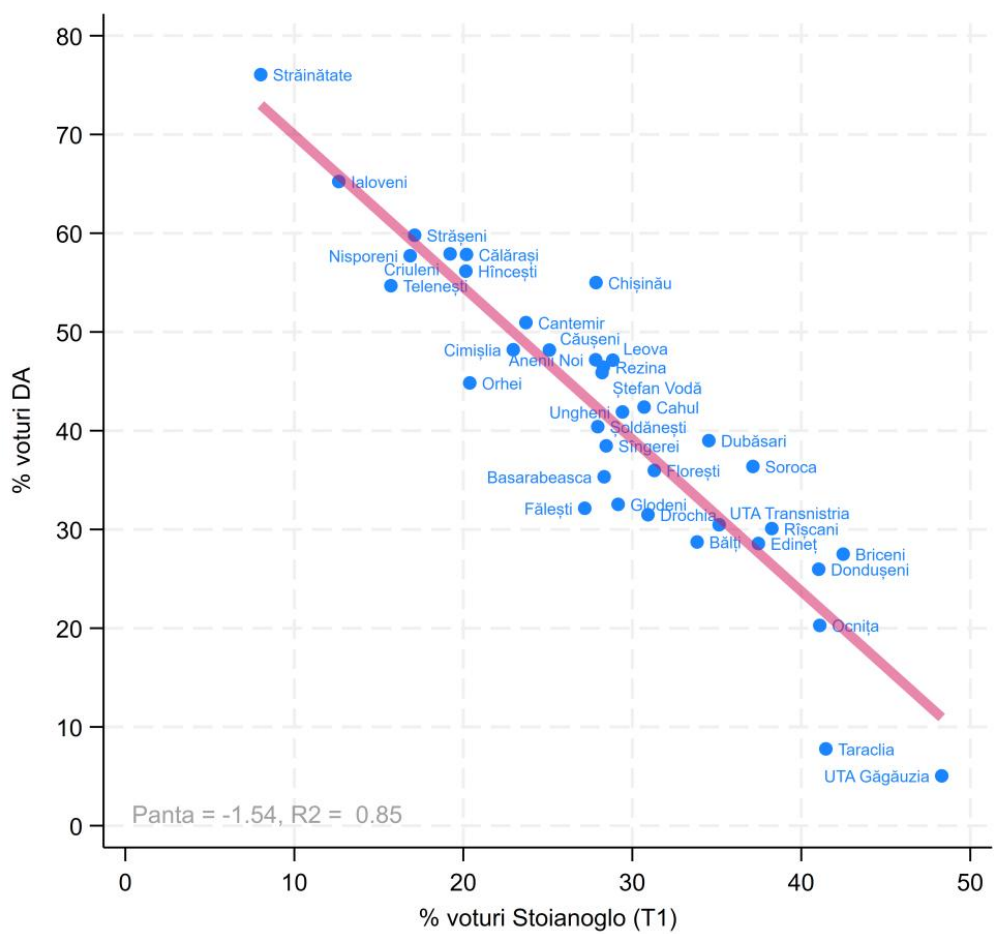
Data: 2024-11-01

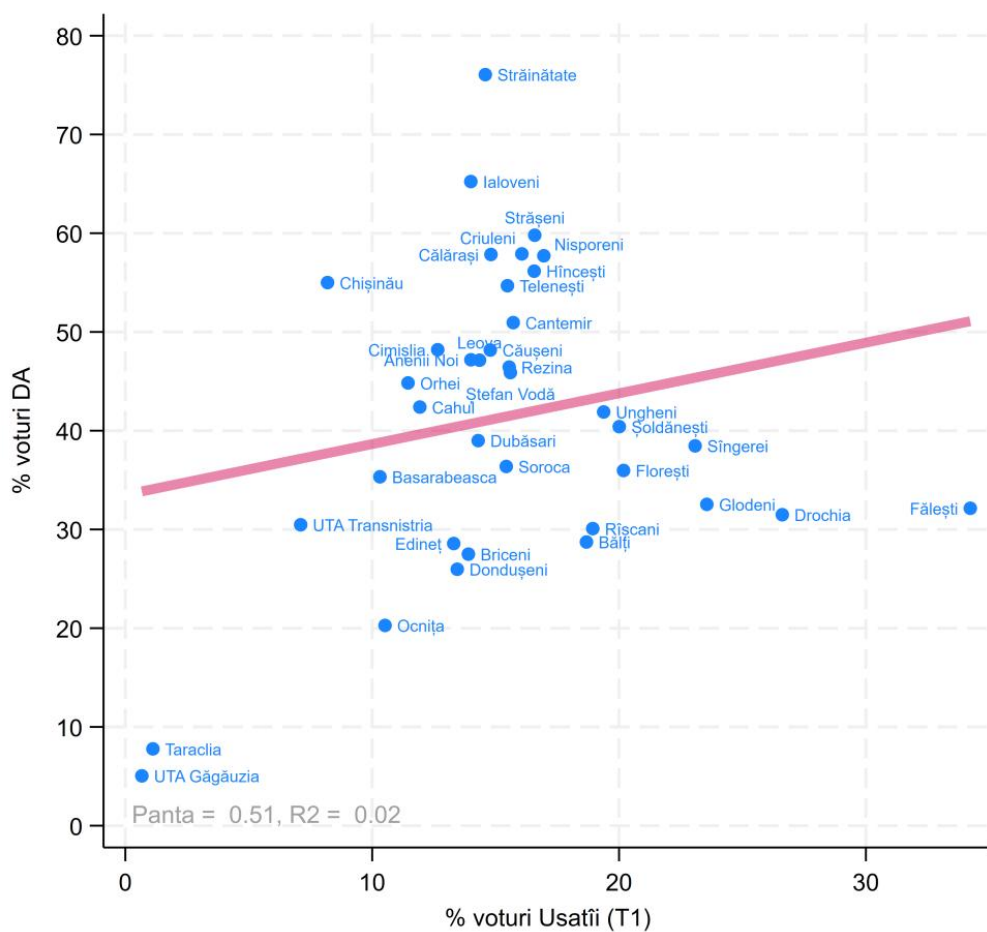
Cuvinte cheie: alegerile-prezidentiale-2024-moldova, referendum-2024-moldova

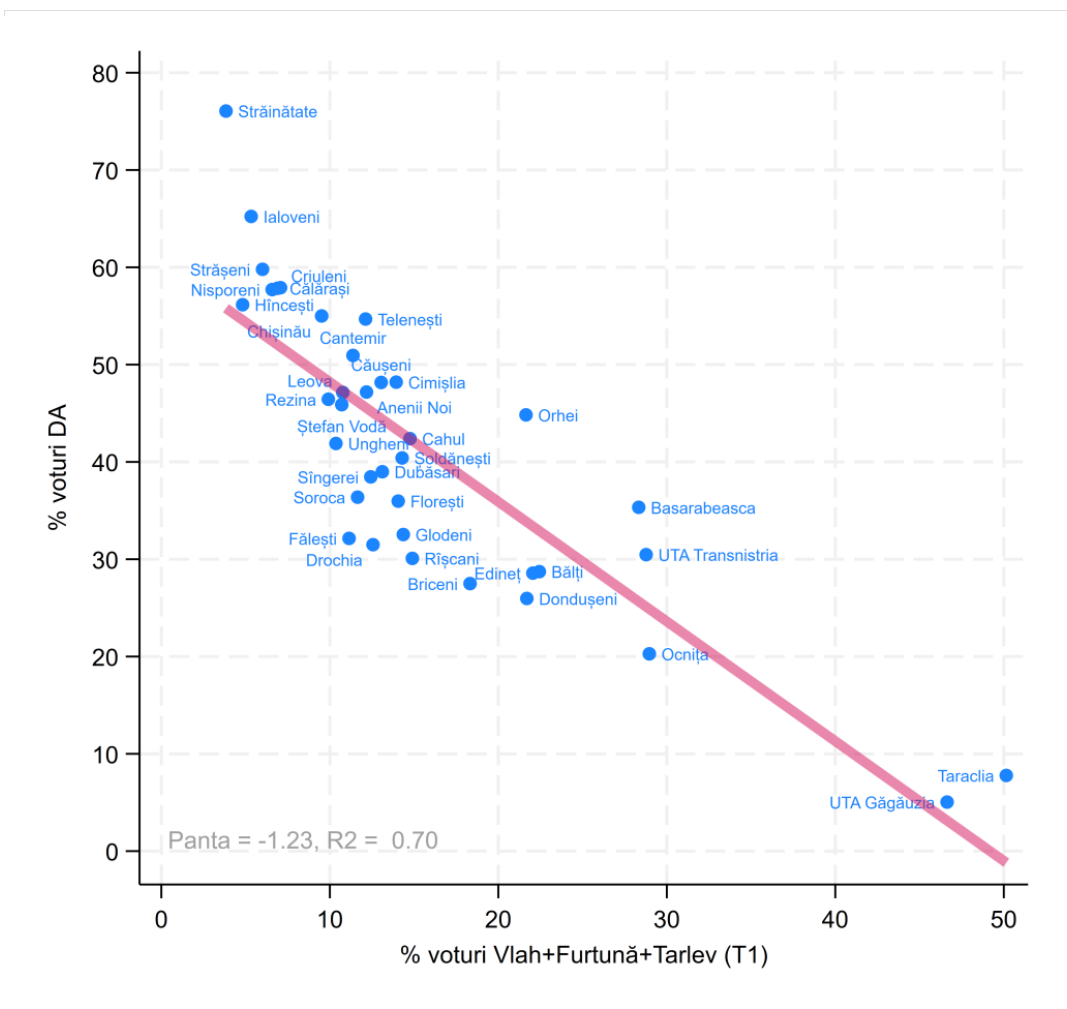
În cazul în care cineva mai avea dubii cu privire la relația dintre ponderea voturilor DA la referendum și ponderea voturilor obținute de candidații la alegerile prezidențiale din Republica Moldova, pun aici patru grafice simple. Mesajul transmis de acestea e cât se poate de clar:

- procentul voturilor DA la referendum corelează pozitiv aproape perfect cu procentul voturilor pro Sandu;
- procentul voturilor DA la referendum corelează puternic negativ cu procentul voturilor pro Stoianoglo;
- procentul voturilor DA la referendum nu corelează cu procentul voturilor pro Usatîi;
- procentul voturilor DA la referendum corelează puternic negativ cu procentele însumate ale candidaților Vlah, Furtună și Tarlev.









Alegeri 2024 Moldova: o analiză la nivel de secție de vot a rezultatelor oficiale

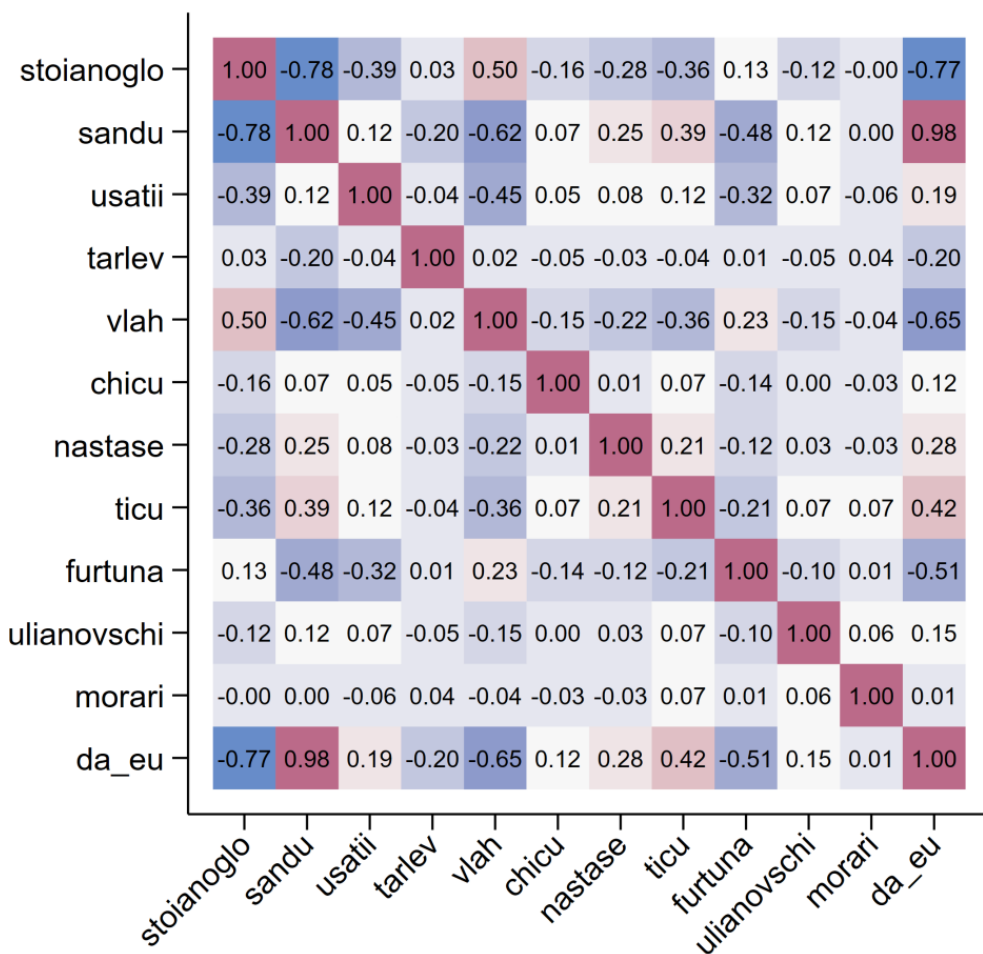
Data: 2024-11-02

Cuvinte cheie: alegerile-prezidentiale-2024-moldova, referendum-2024-moldova

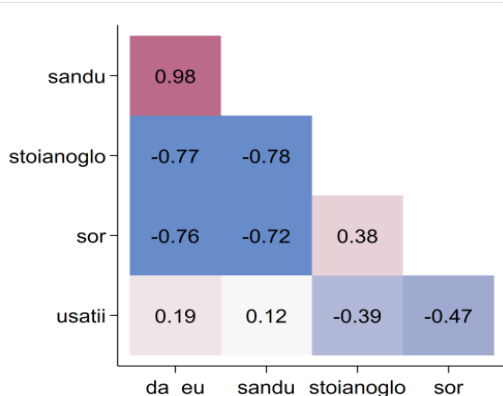
Prezint aici trei tipuri de grafice: (1) corelațiile dintre procente obținute de candidați și procentul DA la referendum, (2) intensitatea votului DA în funcție de procente obținute de diferite perechi de candidați și (3) relația dintre opțiunile de vot pentru candidați și votul DA la referendum pe baza predicțiilor unui model de analiză multivariată. Toate analizele sunt la nivel de secție de vot (doar cele din țară). Câteva concluzii sunt următoarele:

- votul pro Sandu se suprapune perfect cu votul DA;
- votul pro Ticu corelează pozitiv cu votul DA;
- votul pro Usatîi corelează pozitiv puțin cu votul DA;
- votul pro Stoianoglo corelează negativ cu votul DA;
- votul pro Vlah corelează negativ cu votul DA;
- votul pro Furtună corelează negativ cu votul DA;
- votul pro candidați mici susținuți de Șor (suma procentelor obținute de Furtună, Vlah, Talev) corelează negativ cu votul DA;

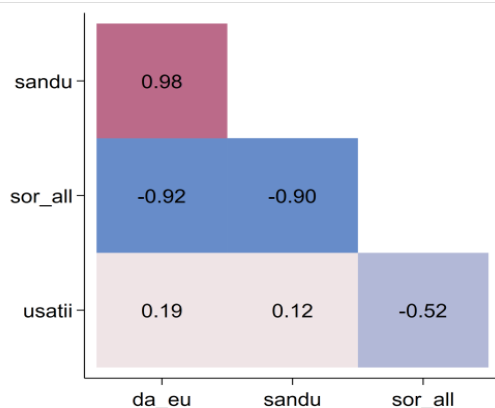
- votul pro candidați susținuți de Șor (suma procentelor obținute de Stoianoglo, Furtună, Vlah, Talev) corelează negativ cu votul DA;
- corelațiile dintre procentele obținute de fiecare dintre candidații susținuți de Șor și votul DA sunt mai mici comparativ cu corelația dintre suma acelorași procente și votul DA;
- votul pro Usatîi corelează ușor pozitiv cu votul pro Sandu, și negativ cu votul pentru candidații susținuți de Șor; din nou, corelația este mai puternică atunci când considerăm suma procentelor obținute de candidații susținuți de Șor;
- modelul predictiv multivariat ne arată că singura relație pozitivă (i-am considerat doar pe primii șase candidați) cu ponderea voturilor DA este în cazul Maiei Sandu; în cazul lui Usatîi relația este ușor negativă; în cazul lui Stoianoglo și al sumei obținute de ceilalți trei candidați susținuți de Șor ambele relații sunt negative; în cazul candidaților susținuți de Șor relația negativă este relativ mai intensă, fiind caracterizată totodată și de o respingere medie relativ mai mare a votului DA.



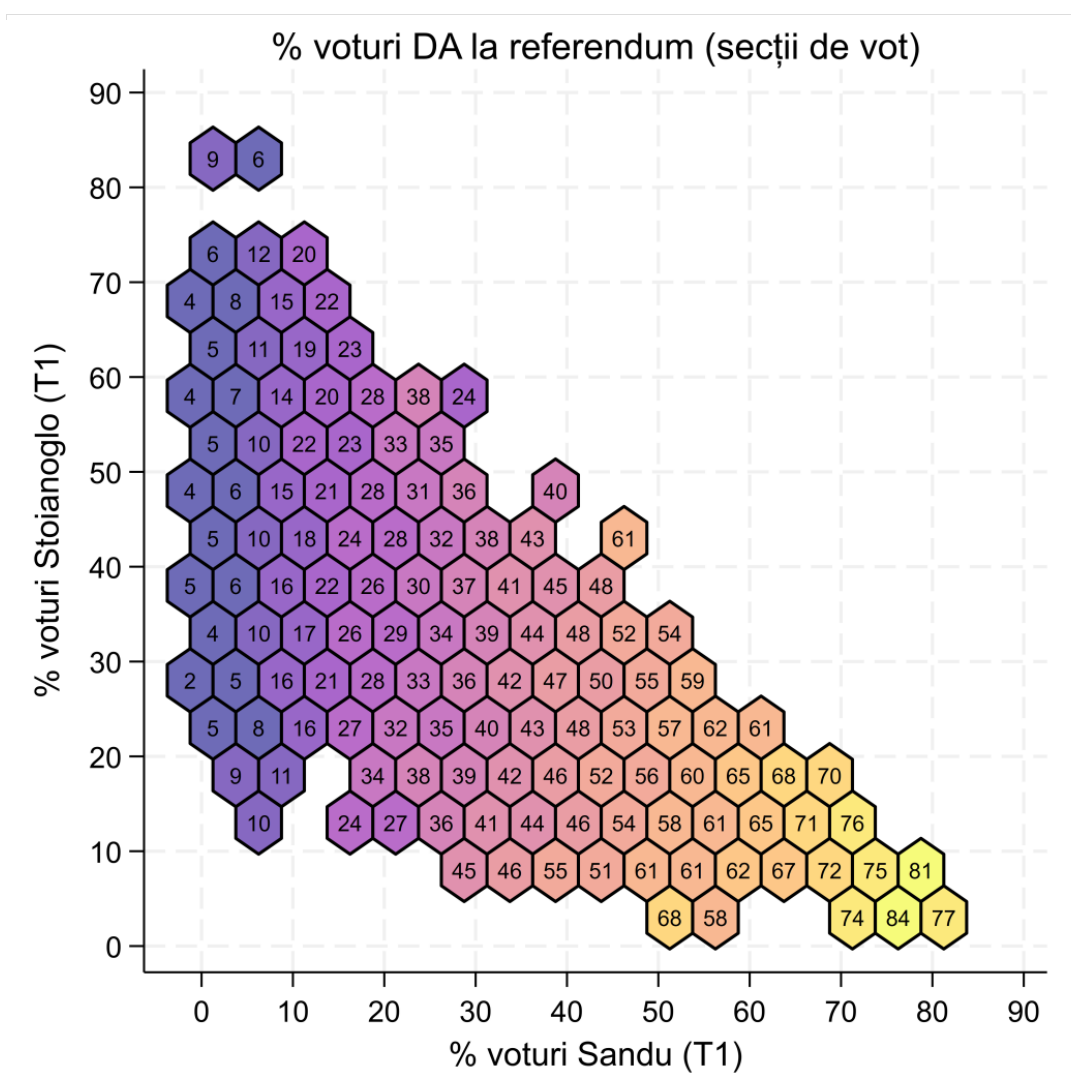
Notă: Matricea de corelație dintre % voturi DA și % voturi candidați la T1 (Moldova, secții din țară).

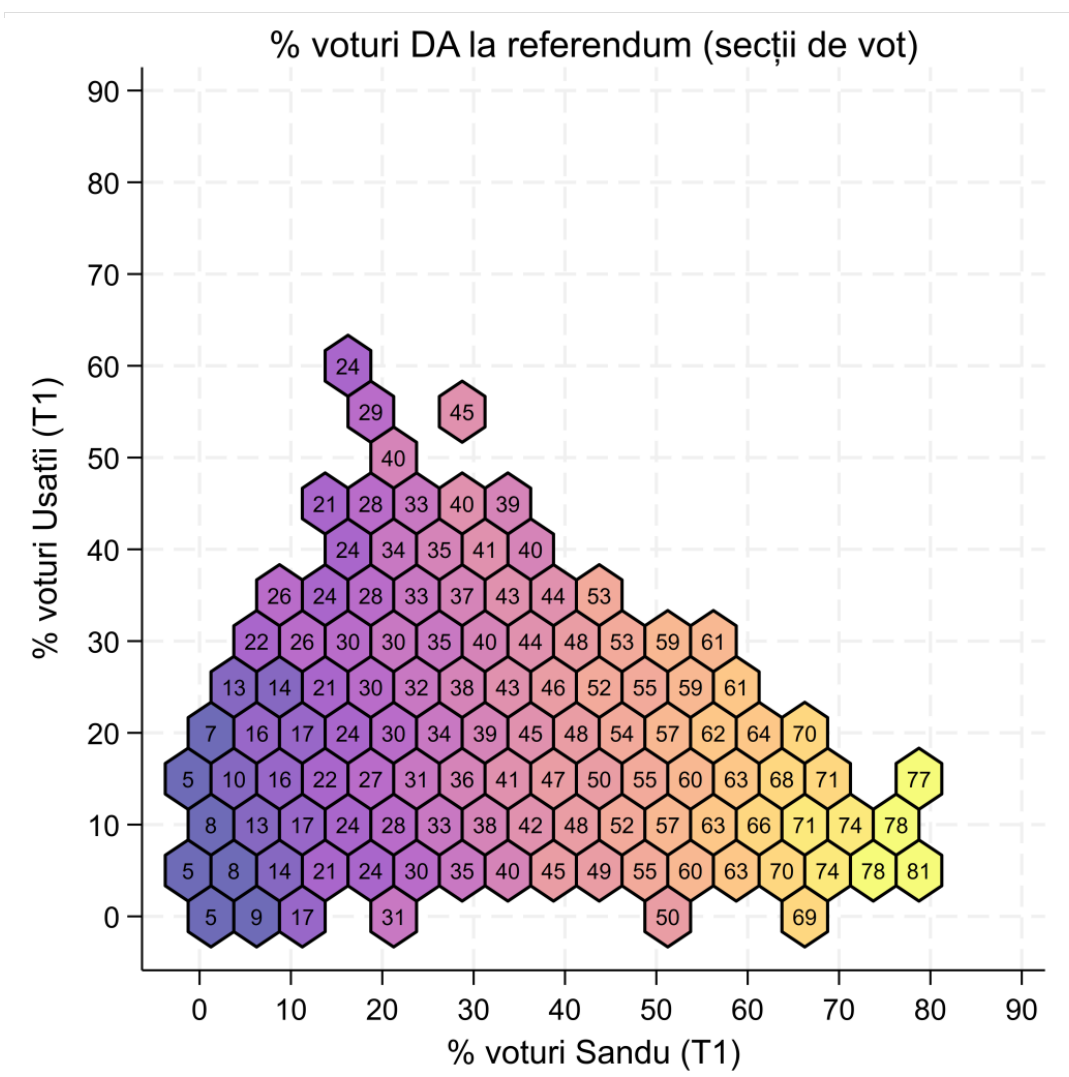


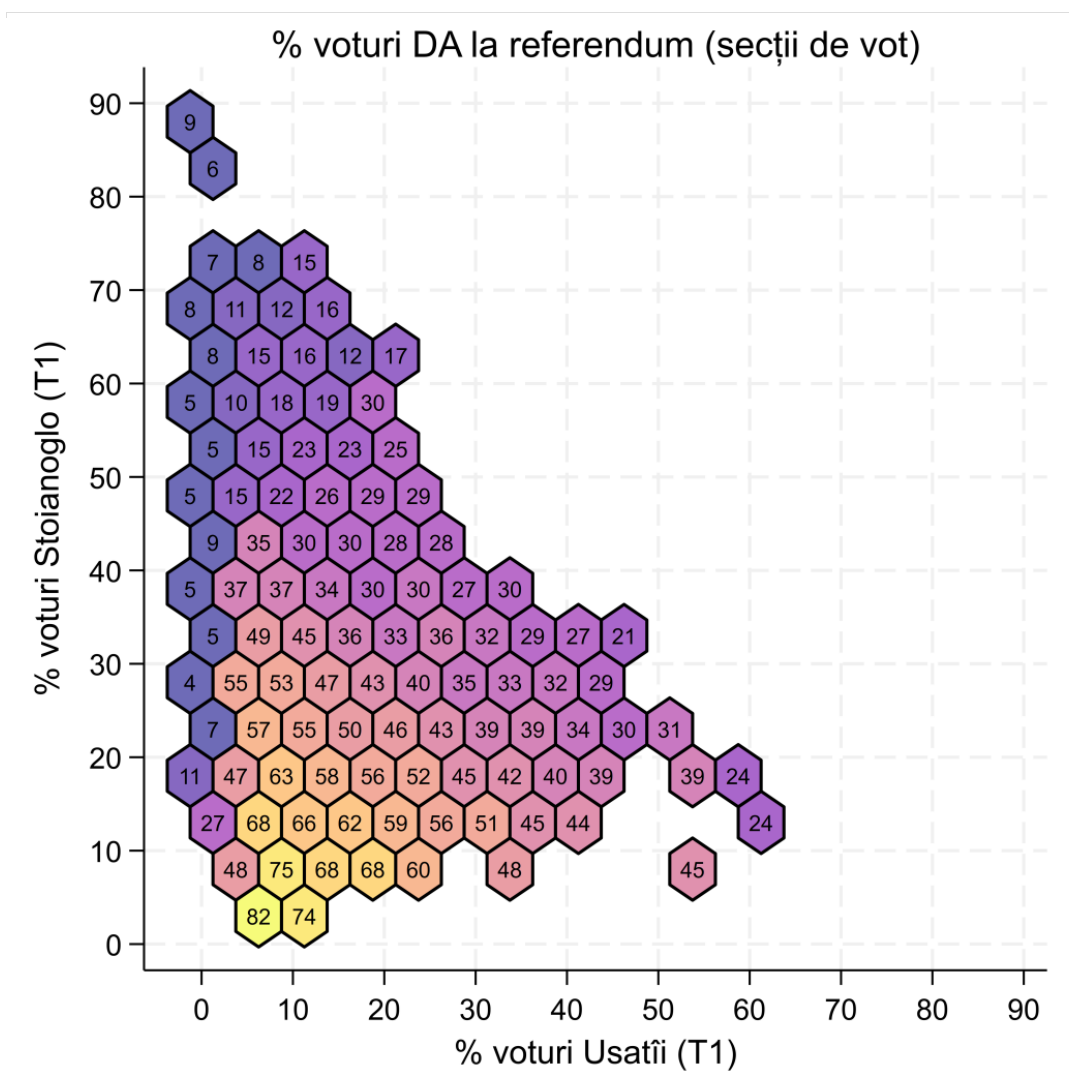
Notă: Matricea de corelație dintre % DA și % candidați la T1 (Moldova, secții din țară).
sor = furtună + vlah + tarlev

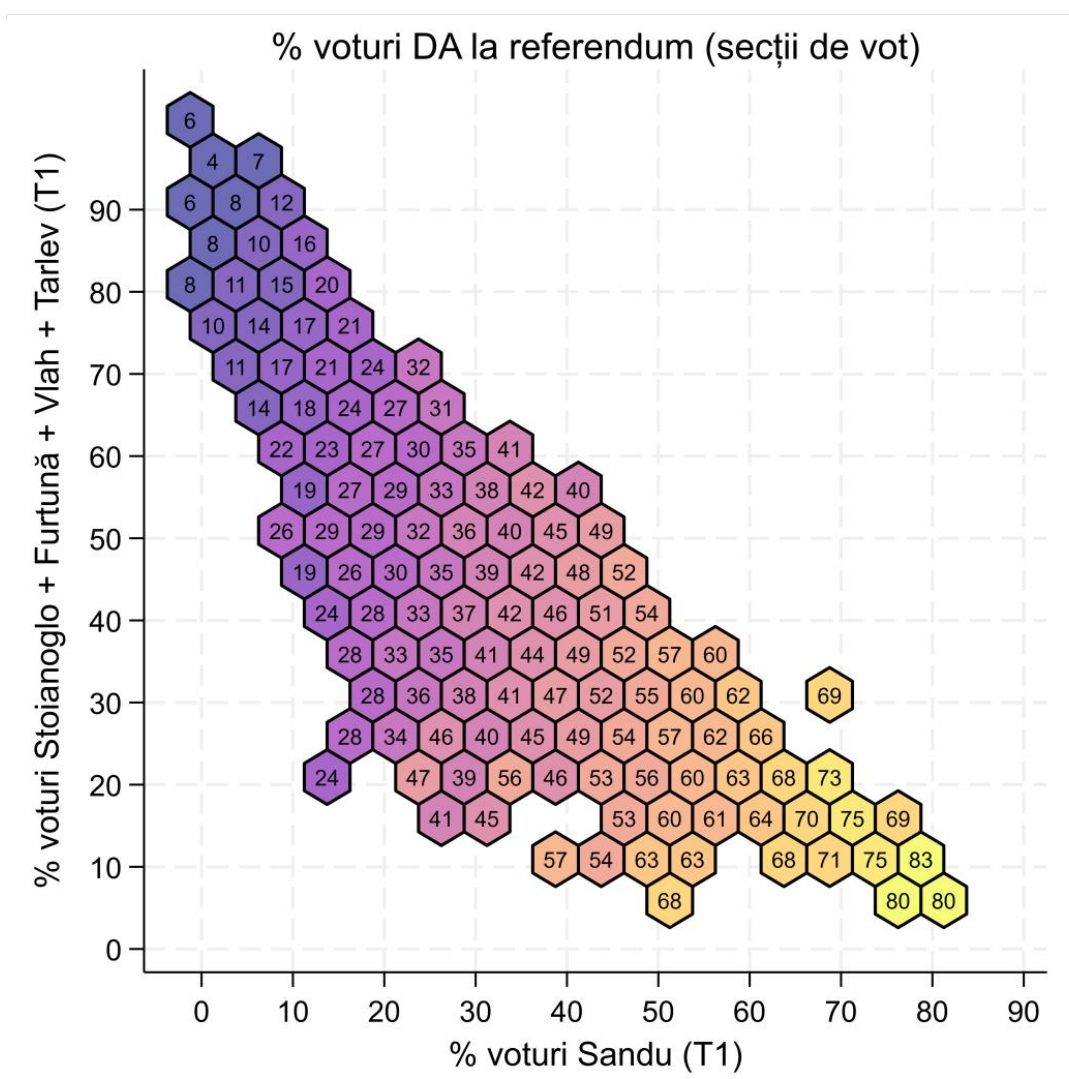


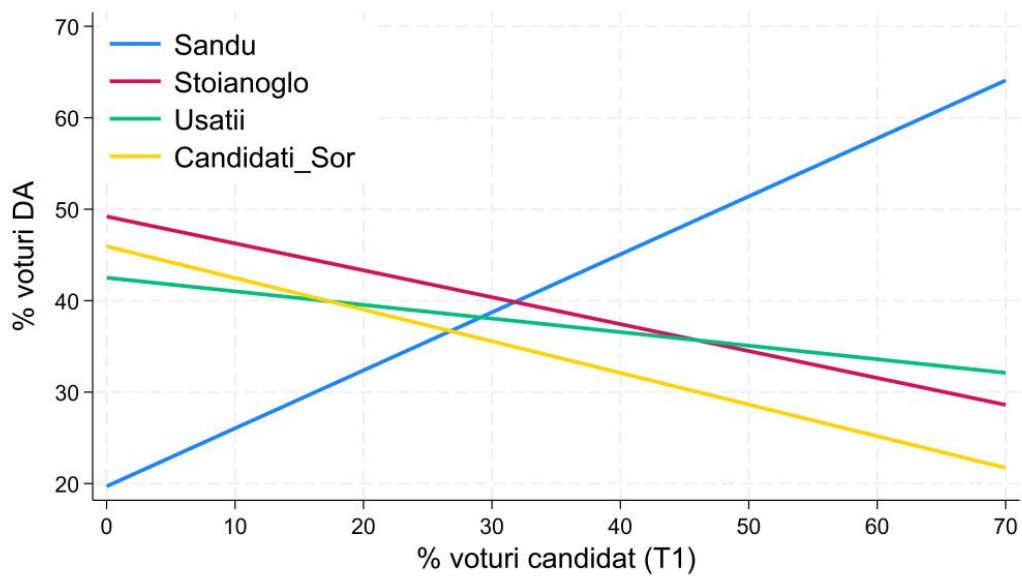
Notă: Matricea de corelație dintre % DA și % candidați la T1 (Moldova, secții din țară).
sor_all = stoianoglo + furtună + vlah + tarlev











Notă: Valori prezise %DA; model tobit; analiză la nivel de secții de vot din țară.

Dinamica intențiilor de vot la alegerile prezidențiale 2024 (1)

Data: 2024-11-08

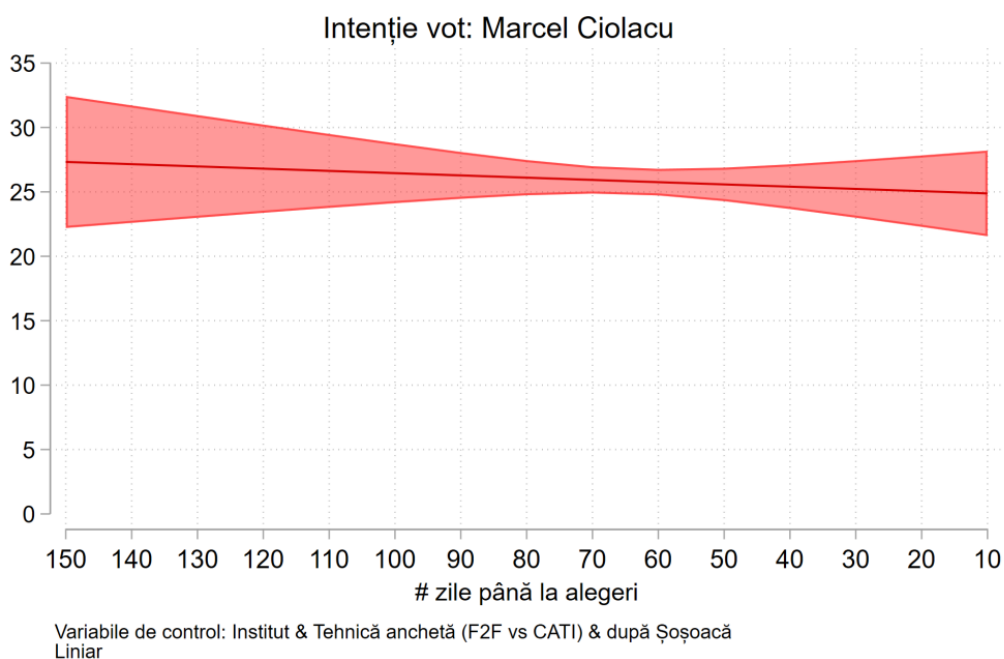
Cuvinte cheie: alegerile-prezidentiale-2024

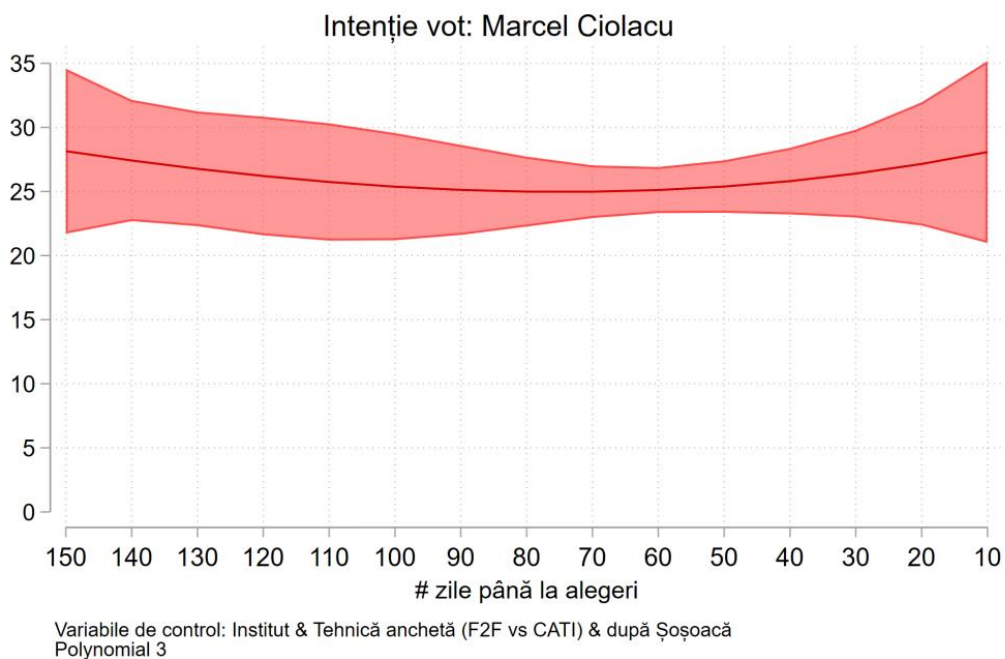
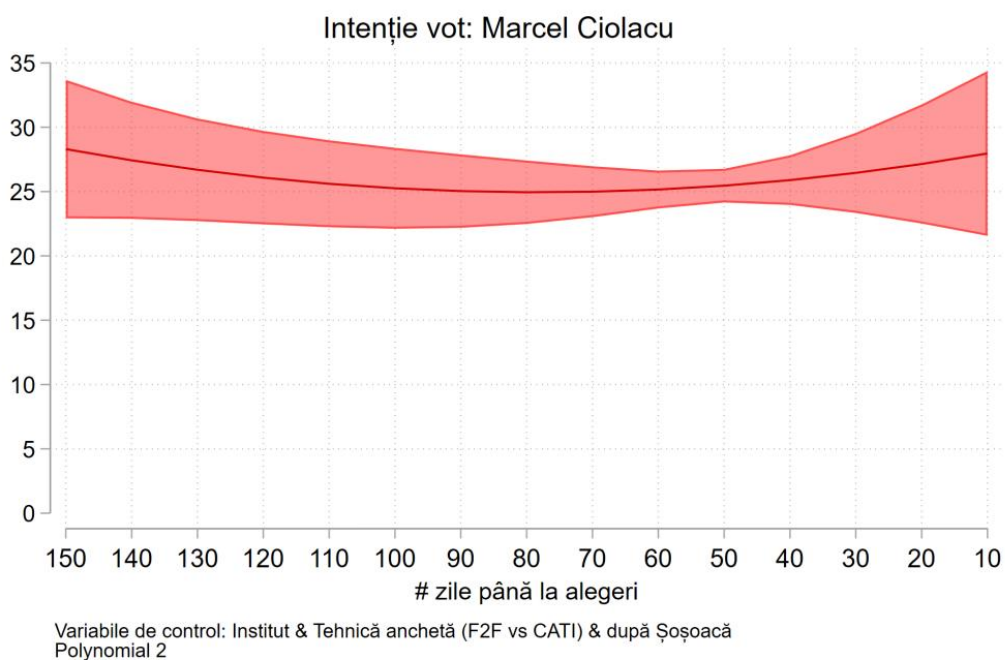
De ceva vreme strâng estimările cu privire la intențiile de vot pentru alegerile prezidențiale așa cum apar în sondajele făcute publice. Cu speranța că aceste grafice vor constitui o sursă de informare și un prilej de reflecție pentru mai multă lume, le postez aici, fără alte comentarii. Important, nu îmi propun să prezint estimările individuale ale sondajelor ci o sinteză a lor pe baza a trei modele statistice. Ultimele sondaje incluse în analiză sunt cele publicate săptămâna aceasta de Sociopol, CURS și Verifield. Din simpatie pentru Nostradamus, am inclus în grafice și tendința indicată de fiecare model până la data de 21 noiembrie (10 zile înainte de alegeri).

Pe scurt, am procedat așa:

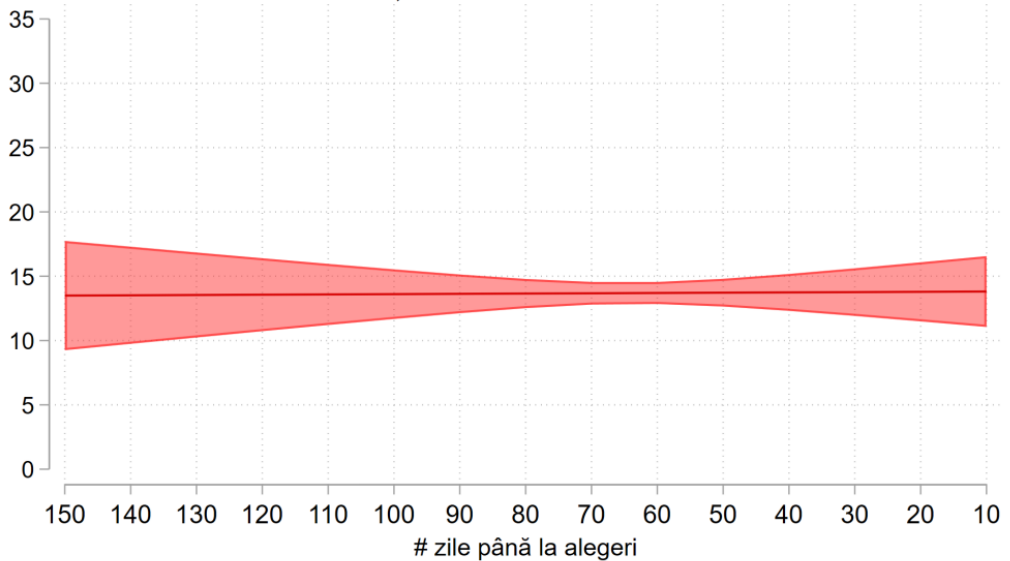
- am considerat toate estimările sondajelor care au fost făcute publice;
- prin estimări mă refer la procente de intenții de vot;
- am estimat trei modele de regresie multiplă multivariată pentru fiecare candidat major (liniar, polynomial 2/3);
- variabila dependentă: intenția de vot estimată de sondaj;

- variabila independentă de interes: numărul de zile până la alegeri;
- variabile de control: institutul care a realizat sondajul, tehnica de anchetă (CATI vs F2F) și eveniment 1 (colectarea datelor s-a realizat înainte sau după decizia CCR de a respinge candidatura Diane Șoșoacă);
- linia roșie reprezintă predicția obținută pe baza modelului statistic;
- banda roșie reprezintă intervalul probabil al valorilor prezise (incertitudinea asociată estimării).



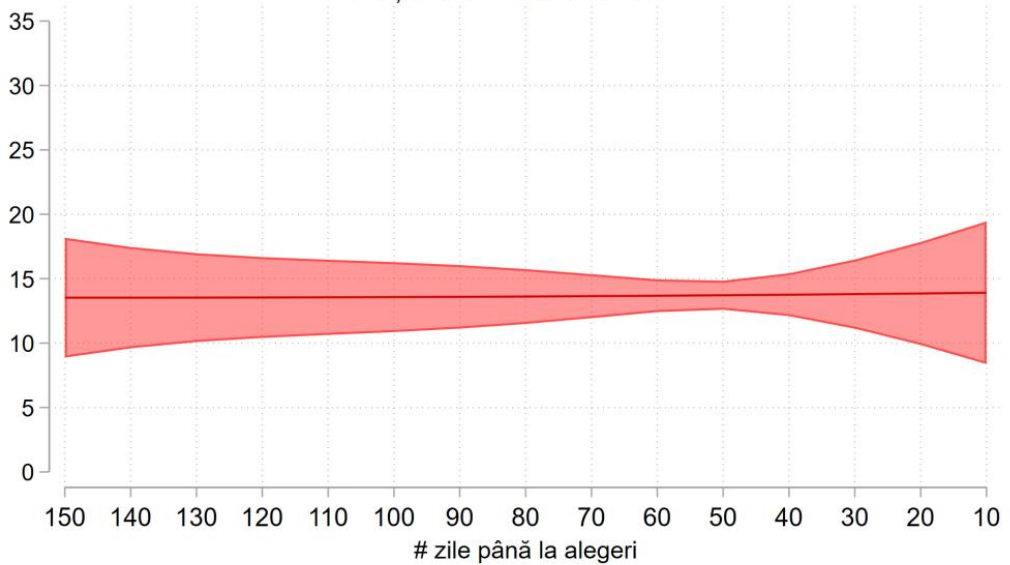


Intenție vot: Nicolae Ciucă



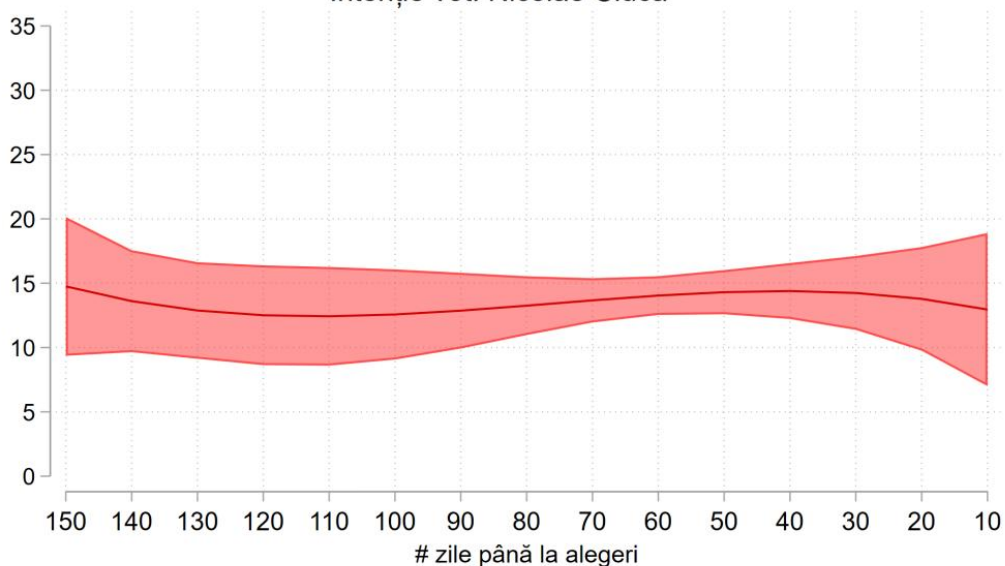
Variabile de control: Institut & Tehnică anchetă (F2F vs CATI) & după Șoșoacă
Liniar

Intenție vot: Nicolae Ciucă



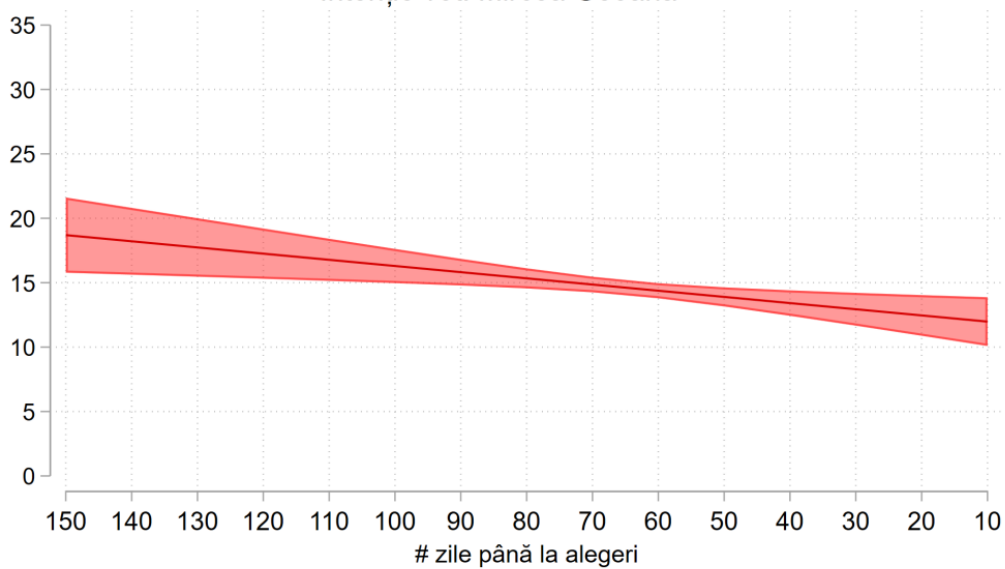
Variabile de control: Institut & Tehnică anchetă (F2F vs CATI) & după Șoșoacă
Polynomial 2

Intenție vot: Nicolae Ciucă



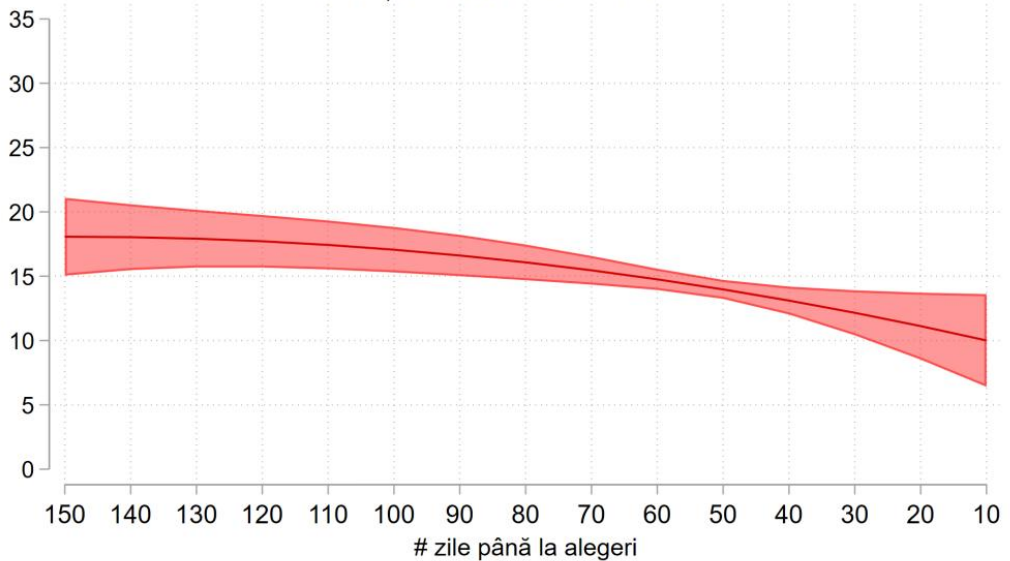
Variabile de control: Institut & Tehnică anchetă (F2F vs CATI) & după Șoșoacă
Polynomial 3

Intenție vot: Mircea Geoană



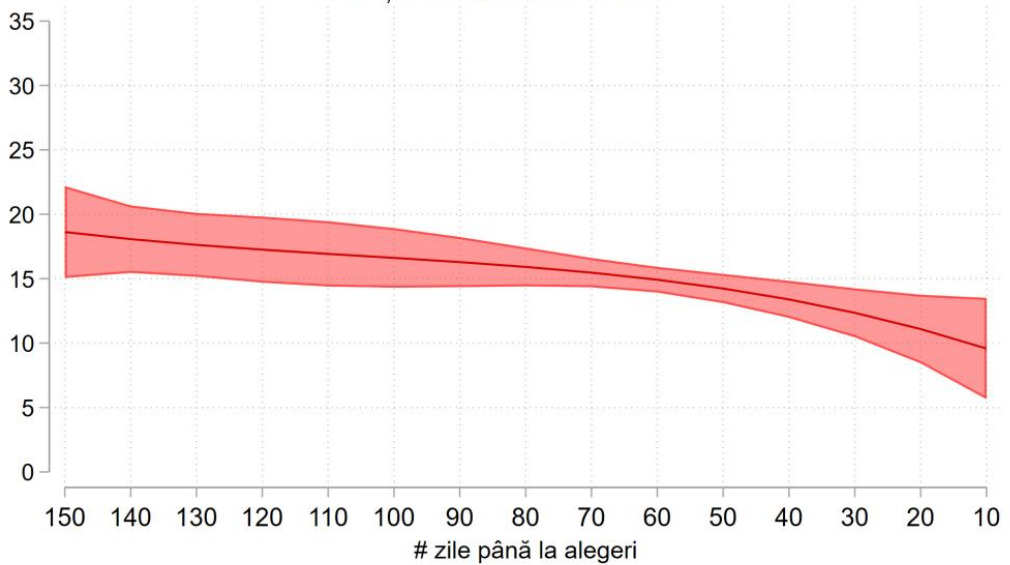
Variabile de control: Institut & Tehnică anchetă (F2F vs CATI) & după Șoșoacă
Liniar

Intenție vot: Mircea Geoană



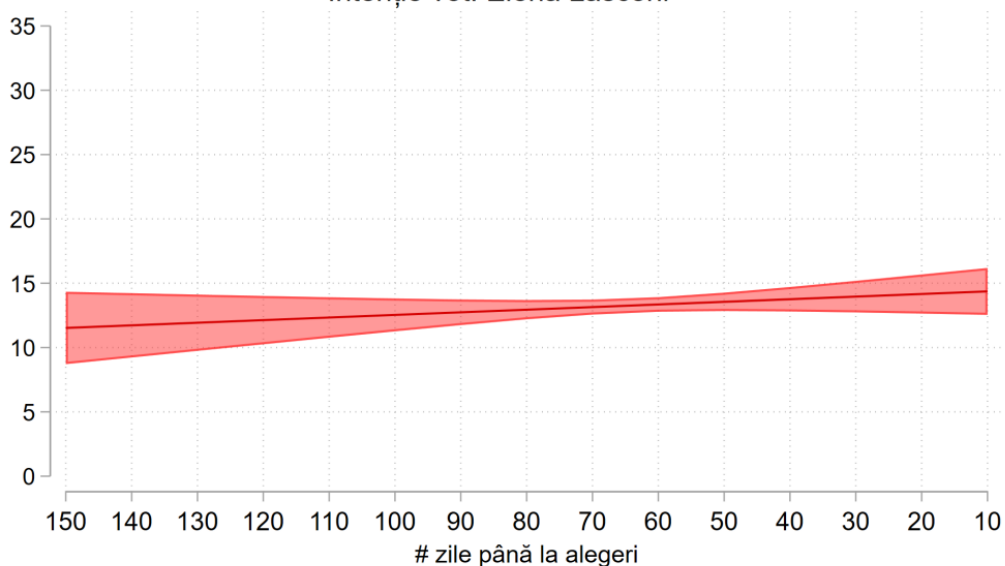
Variabile de control: Institut & Tehnică anchetă (F2F vs CATI) & după Șoșoacă
Polynomial 2

Intenție vot: Mircea Geoană



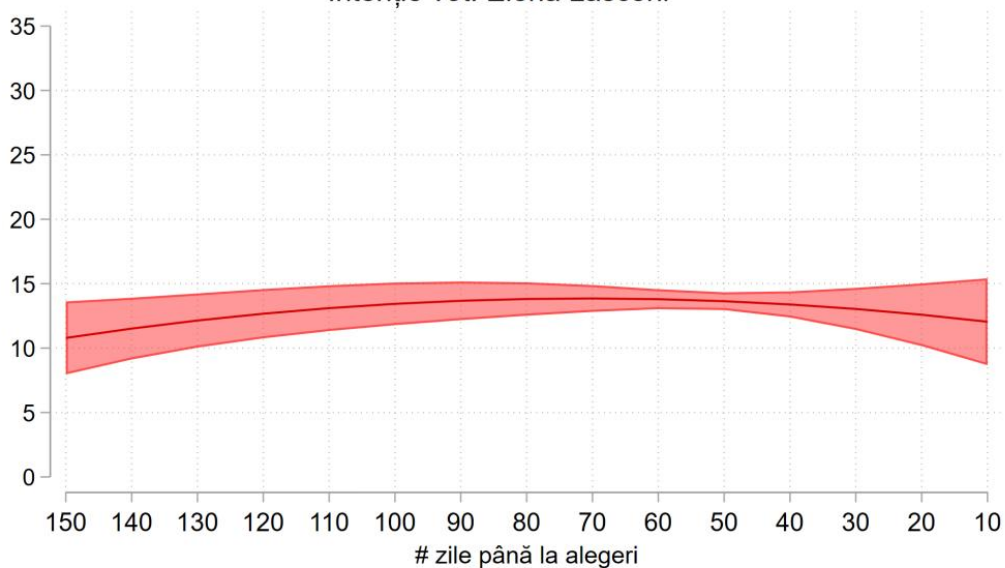
Variabile de control: Institut & Tehnică anchetă (F2F vs CATI) & după Șoșoacă
Polynomial 3

Intenție vot: Elena Lasconi



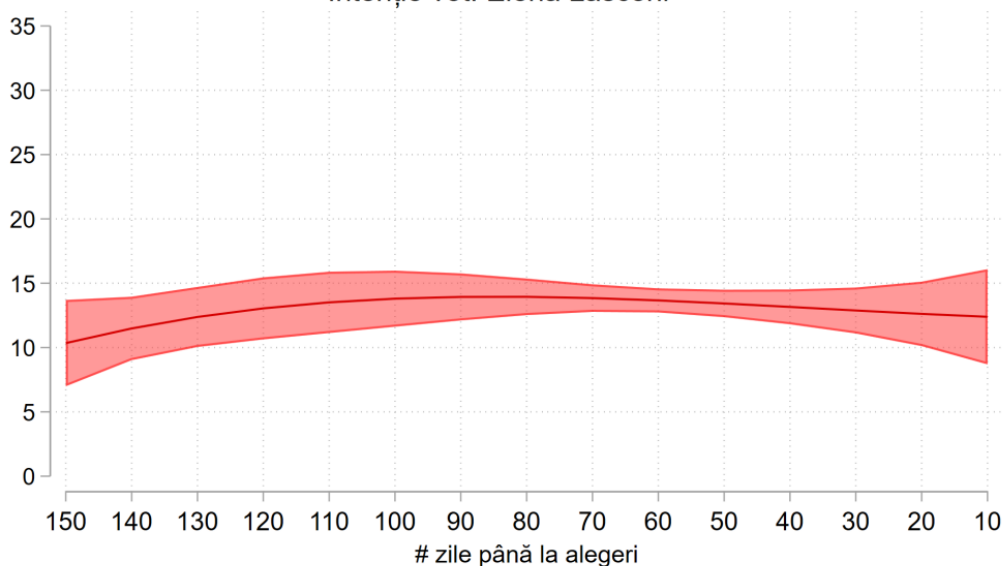
Variabile de control: Institut & Tehnică anchetă (F2F vs CATI) & după Șoșoacă
Liniar

Intenție vot: Elena Lasconi



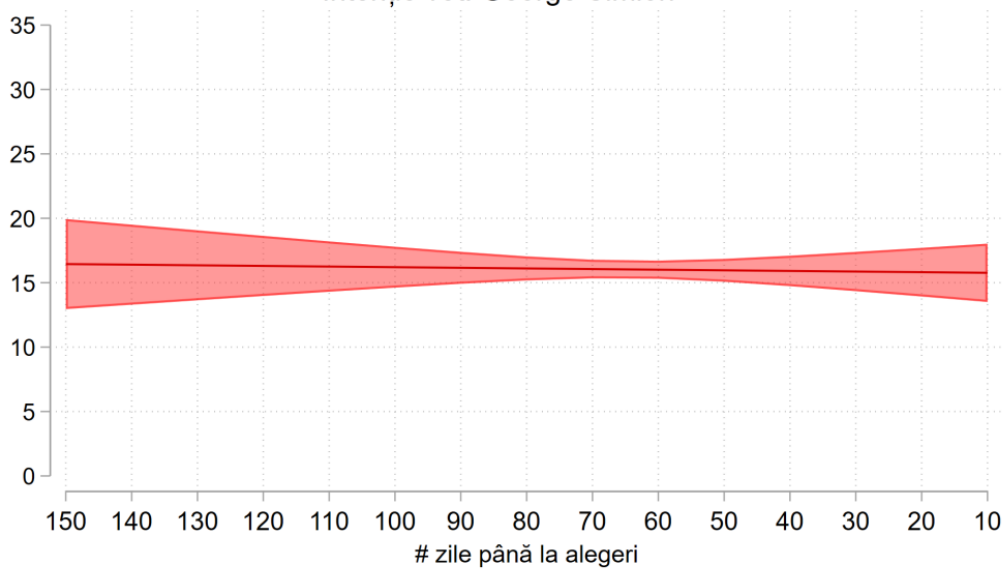
Variabile de control: Institut & Tehnică anchetă (F2F vs CATI) & după Șoșoacă
Polynomial 2

Intenție vot: Elena Lasconi

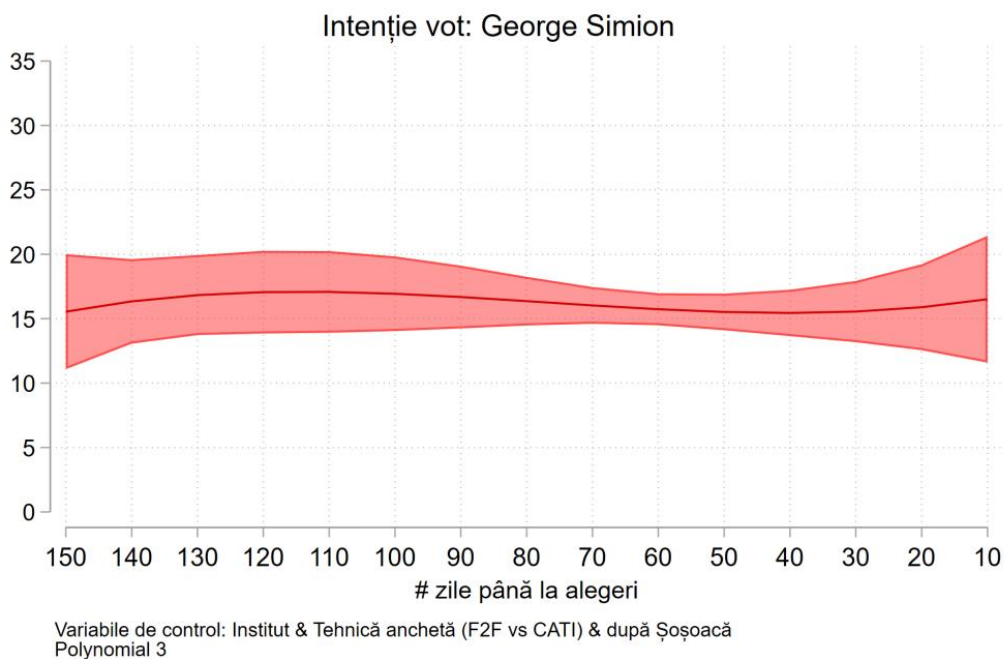
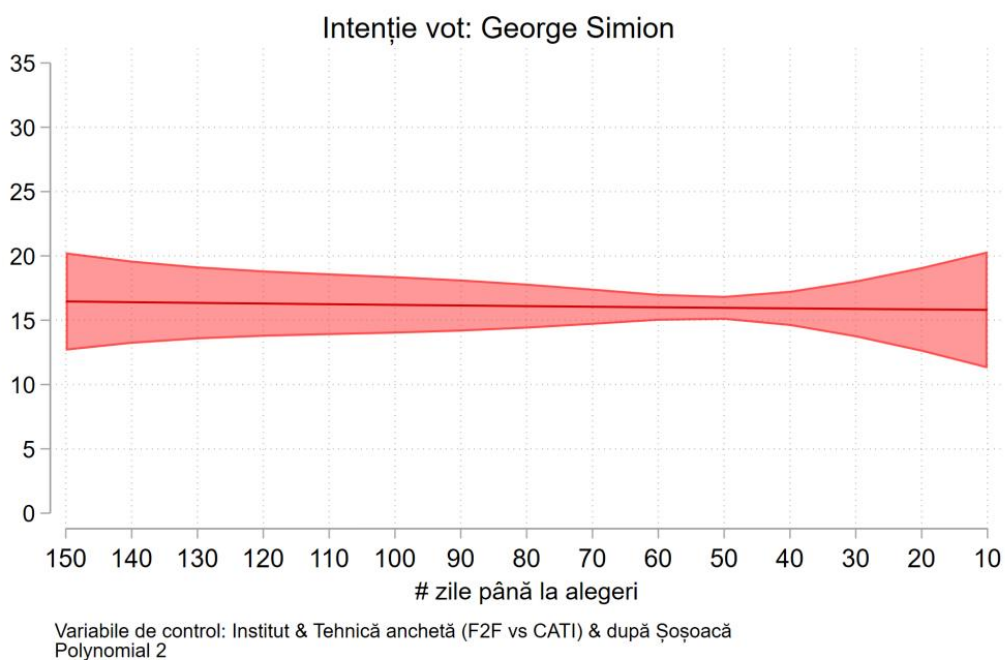


Variabile de control: Institut & Tehnică anchetă (F2F vs CATI) & după Șoșoacă
Polynomial 3

Intenție vot: George Simion



Variabile de control: Institut & Tehnică anchetă (F2F vs CATI) & după Șoșoacă
Liniar



„Eu cu cine votez?”

Data: 2024-11-08

Cuvinte cheie: alegerile-parlamentare-2024

Ieri a fost zarvă mare. S-a zvonit în oraș că mai mulți alegători turmentați au ieșit în stradă cerând în gura mare să le spună cineva cu cine să voteze! Cum e departe de mine gândul că aș fi competent să răspund la astfel de întrebări (nu că mi-aș dori acest lucru), i-am sfătuit să vadă care sunt preferințele partidelor în relație cu preferințele lor. Pentru a primi un posibil răspuns, e suficient să viziteze aplicația de la adresa <https://eucucinevotez.eu/>. Nu costă nimic și nici nu doare. După un diagnostic rapid în 28 de pași, aplicația vă indică compatibilitatea cu o serie de partide. Desigur, ceea ce va decide fiecare alegător în urma recomandărilor primite rămâne strict secret.

Disclaimer: Nu suntem responsabili pentru eventualele șocuri provocate de rezultatul consultației.

Dinamica intențiilor de vot la alegerile parlamentare 2024 (1)

Data: 2024-11-09

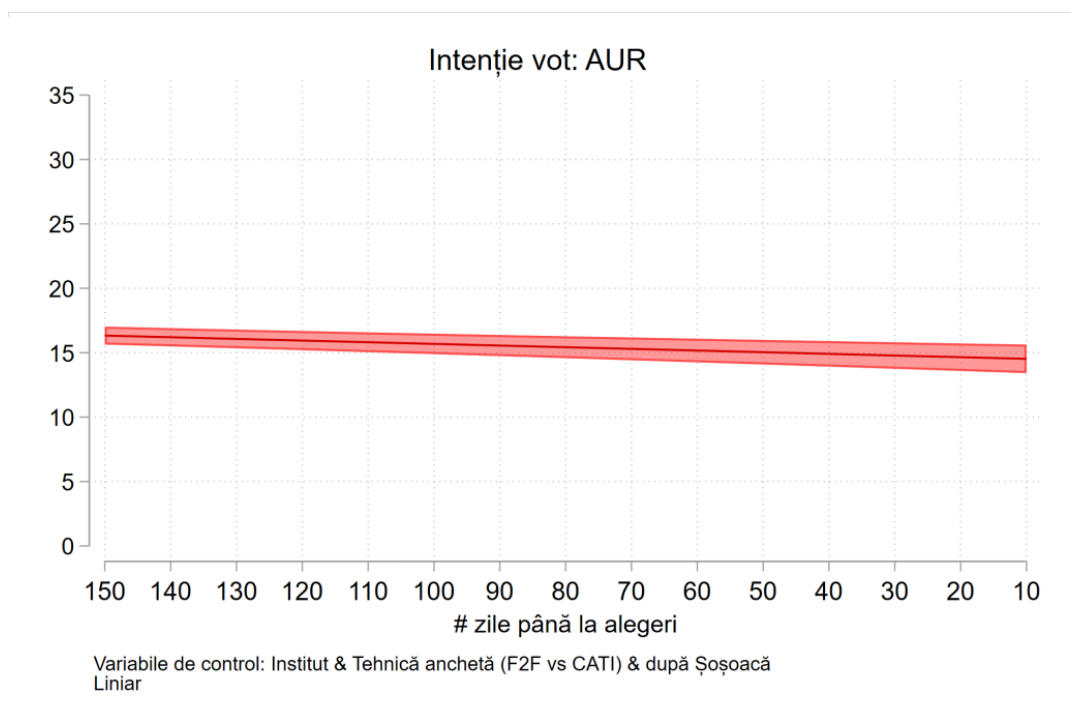
Cuvinte cheie: alegerile-parlamentare-2024

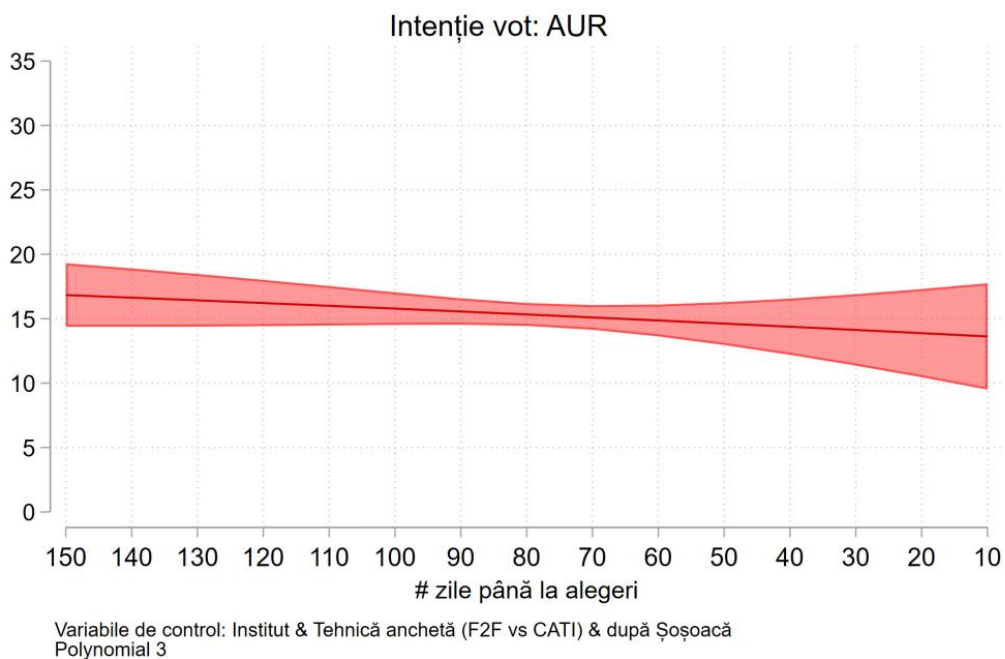
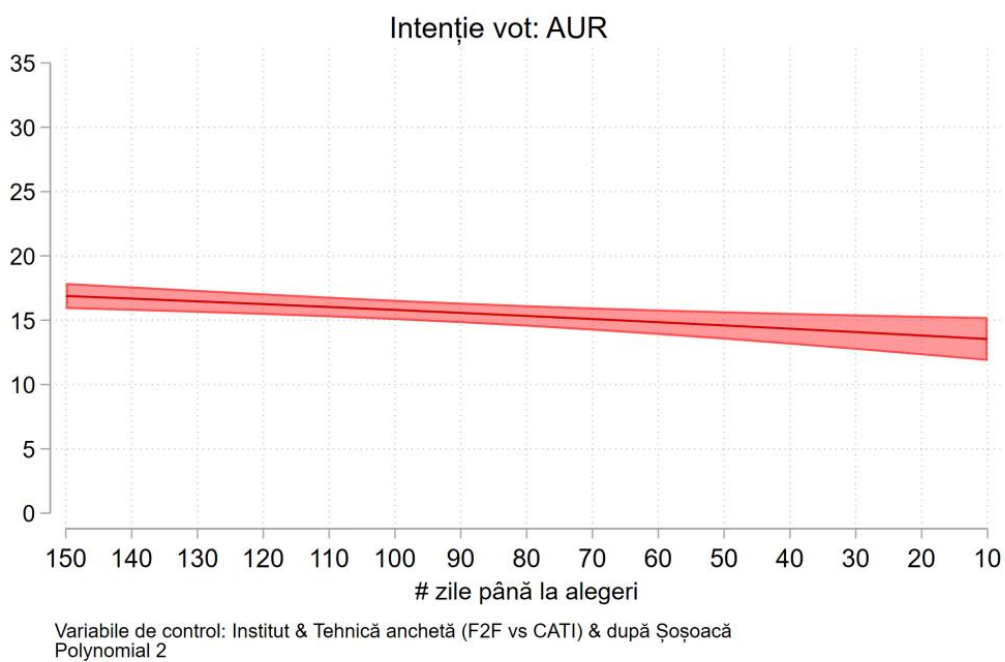
De ceva vreme strâng estimările cu privire la intențiile de vot pentru alegerile parlamentare așa cum apar în sondajele făcute publice. Cu speranța că aceste grafice vor constitui o sursă de informare și un prilej de reflecție pentru mai multă lume, le postez aici, fără alte comentarii. Important, nu îmi propun să prezint estimările individuale ale sondajelor ci o sinteză a lor pe baza a trei modele statistice. Ultimele sondaje incluse în analiză sunt cele publicate săptămâna aceasta de Sociopol, CURS și Verifield. Din simpatie pentru Nostradamus, am inclus în grafice și tendința indicată de fiecare model până la data de 28 noiembrie (10 zile înainte de alegeri).

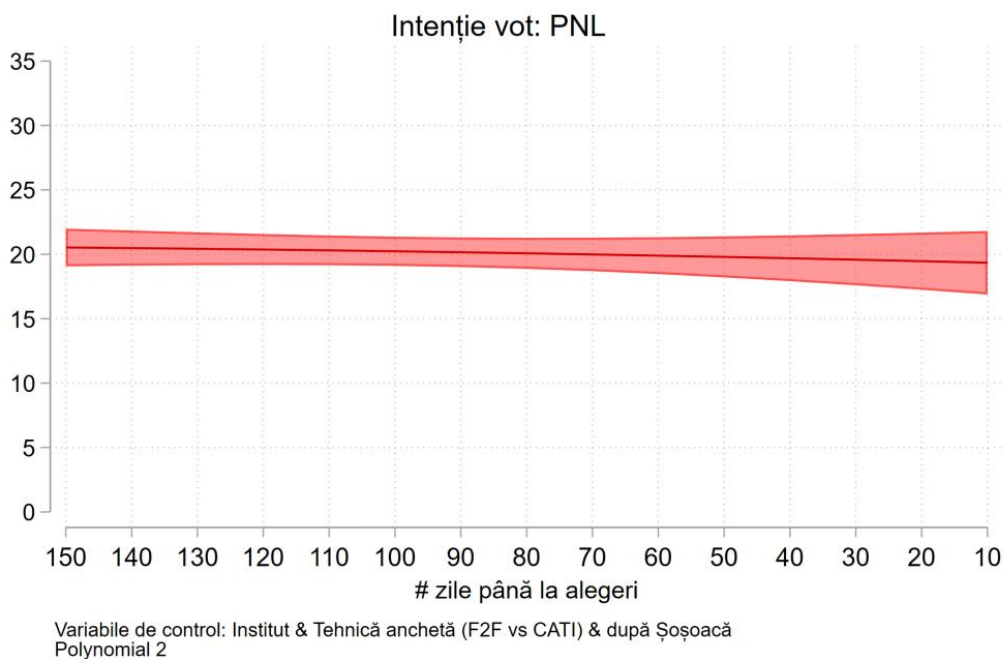
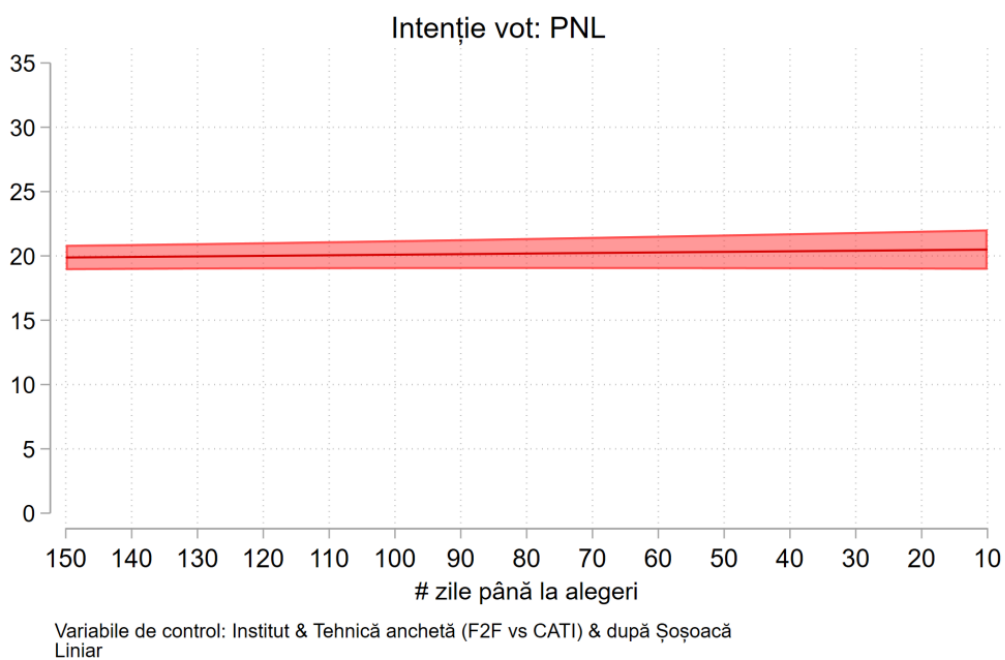
Pe scurt, am procedat așa:

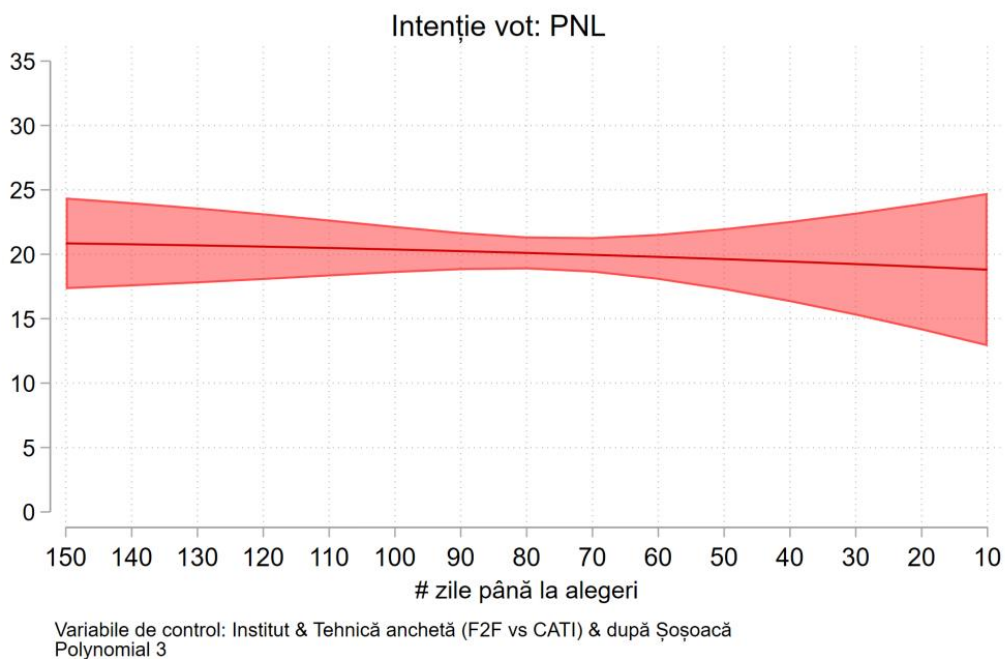
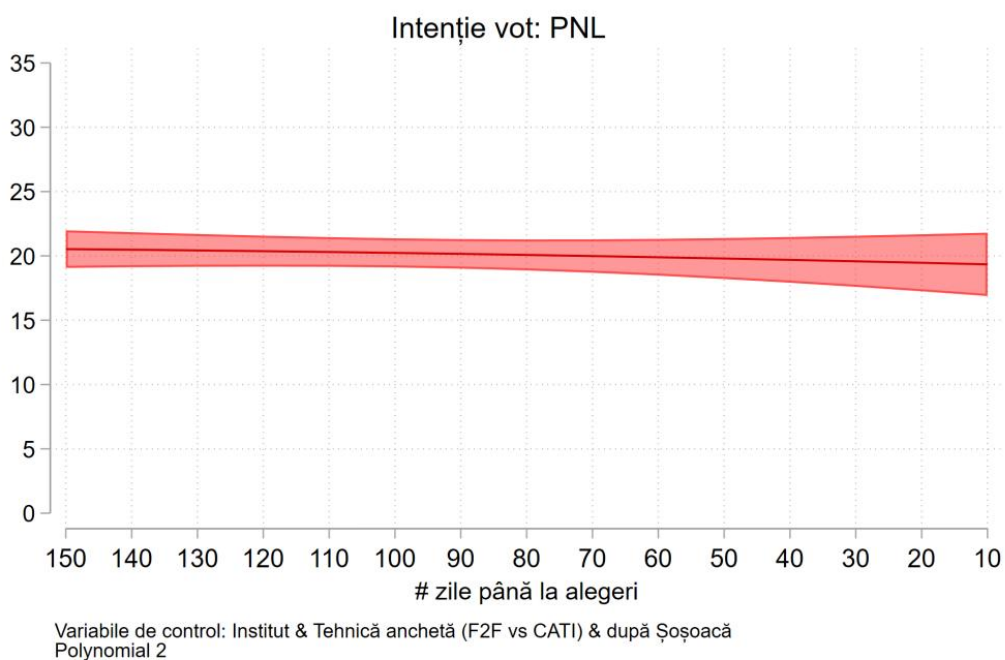
- am considerat toate estimările sondajelor care au fost făcute publice;
- prin estimări mă refer la procente de intenții de vot;
- am estimat trei modele de regresie multiplă multivariată pentru fiecare partid major (liniar, polynomial 2/3);
- variabila dependentă: intenția de vot estimată de sondaj;

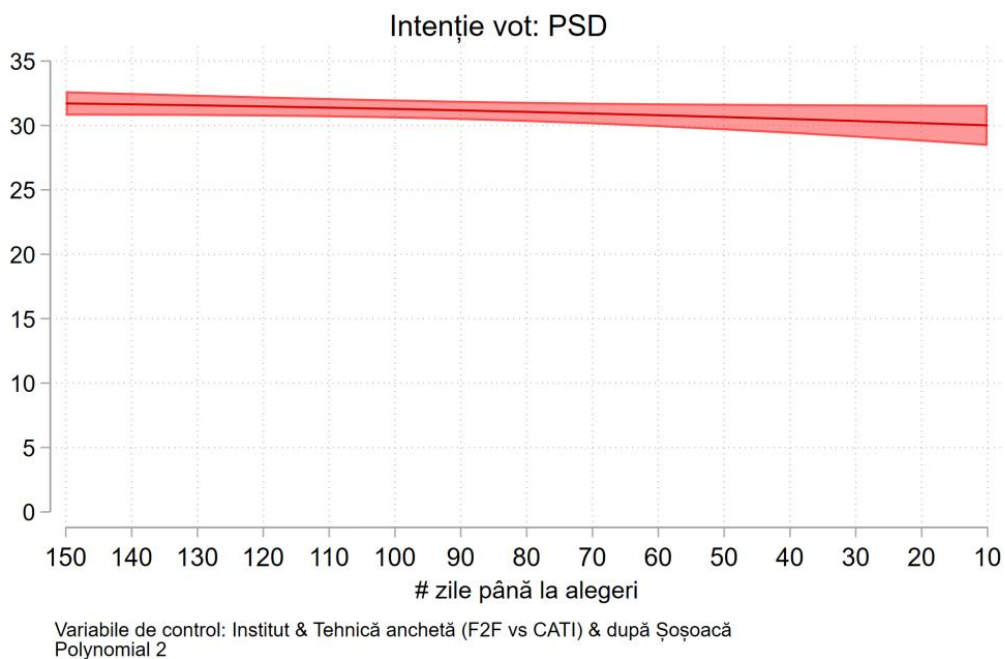
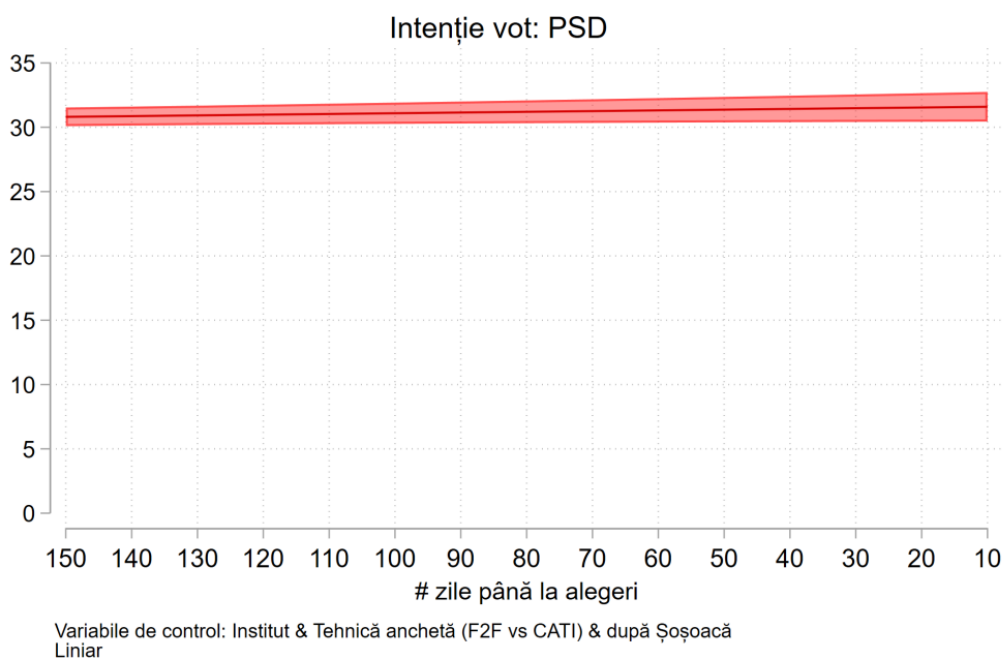
- variabila independentă de interes: numărul de zile până la alegeri;
- variabile de control: institutul care a realizat sondajul, tehnica de anchetă (CATI vs F2F) și eveniment 1 (colectarea datelor s-a realizat înainte sau după decizia CCR de a respinge candidatura Diane Șoșoacă);
- linia roșie reprezintă predicția obținută pe baza modelului statistic;
- banda roșie reprezintă intervalul probabil al valorilor prezise (incertitudinea asociată estimării).

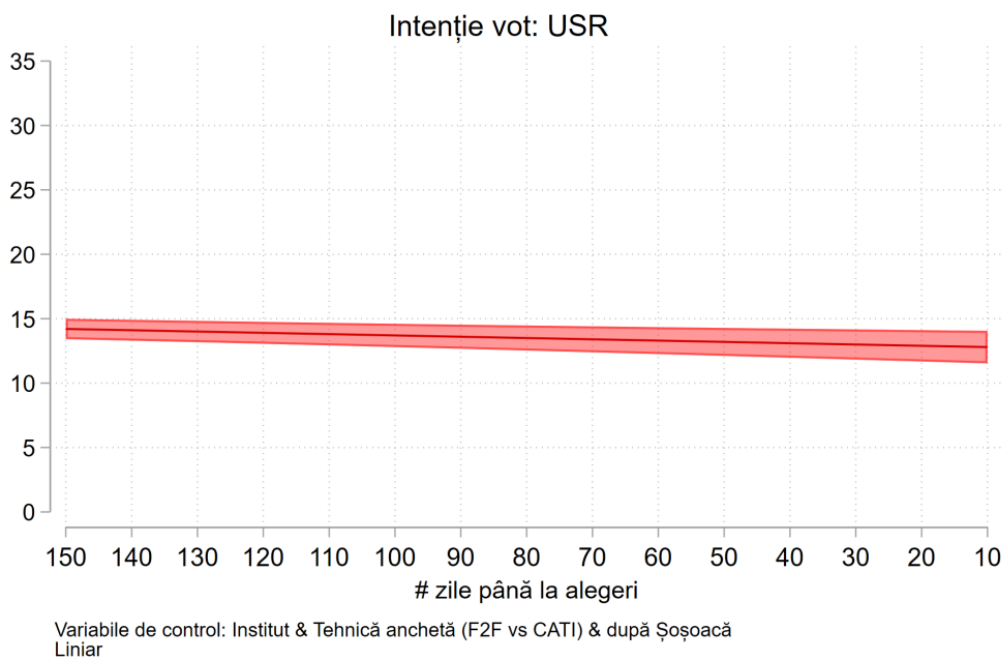
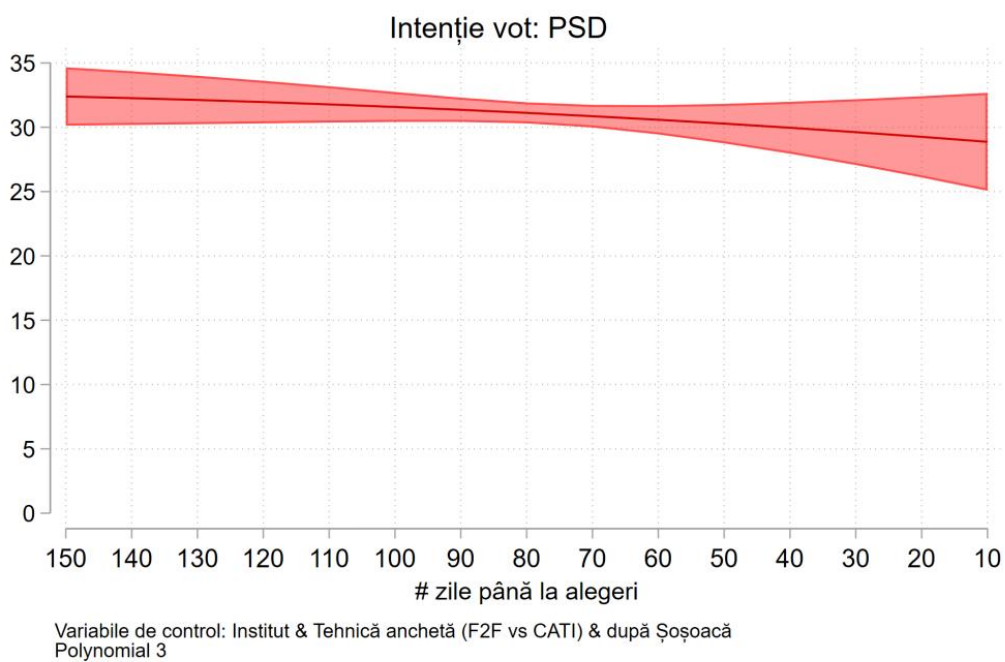


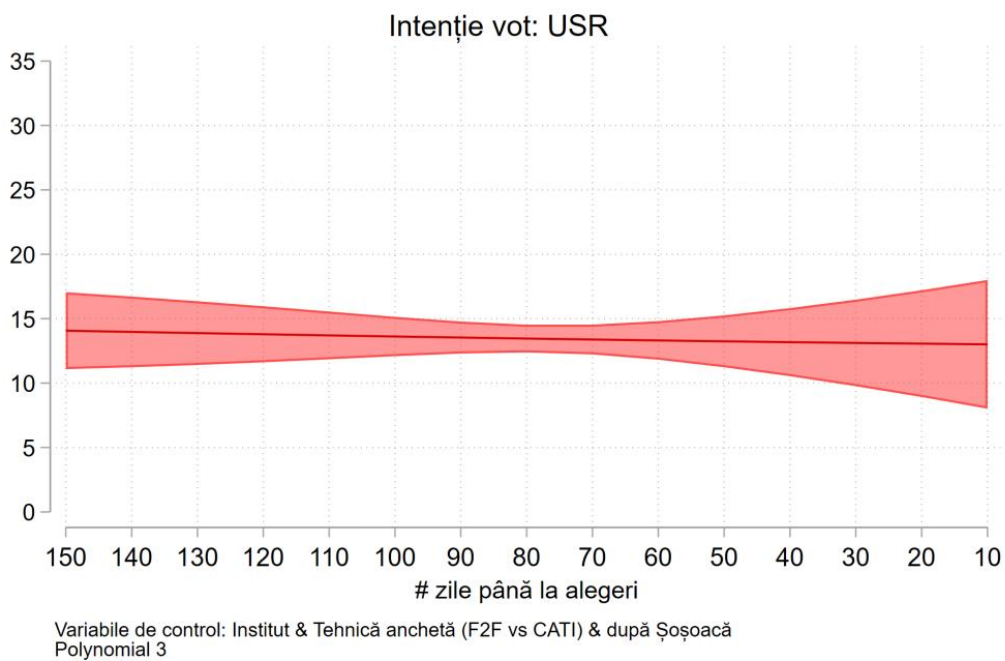
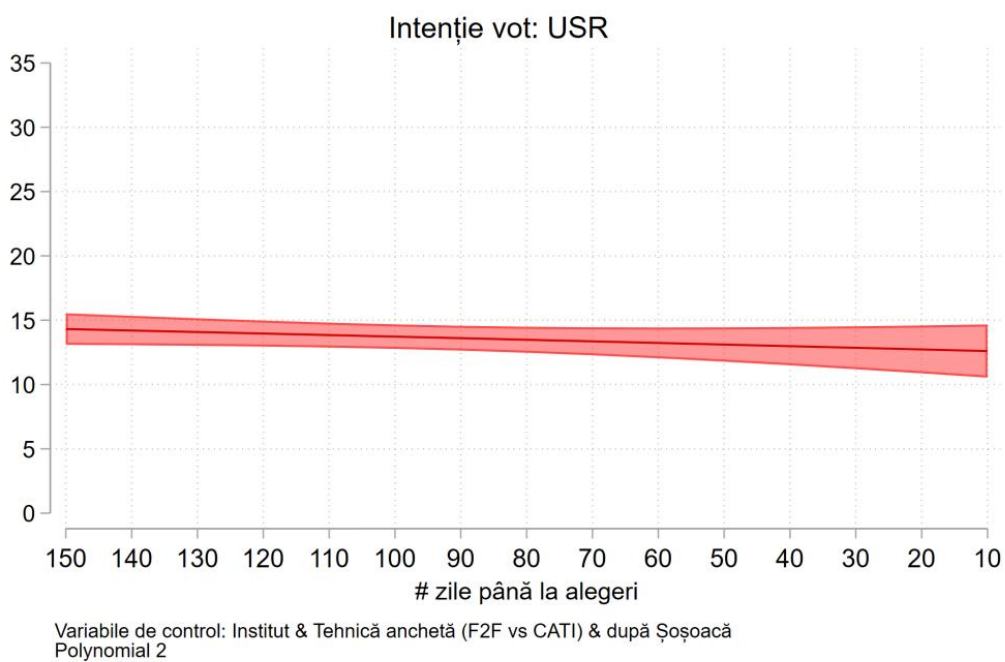












„House effects”: alegerile prezidențiale 2024

Data: 2024-11-10

Cuvinte cheie: alegerile-prezidentiale-2024, house-effects

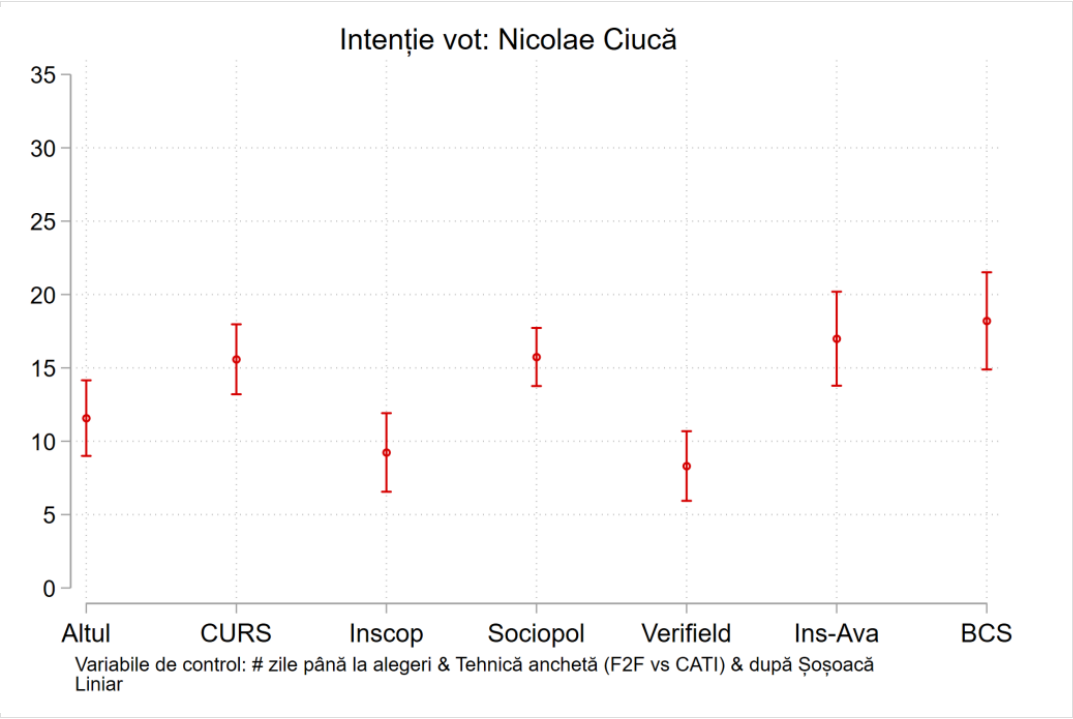
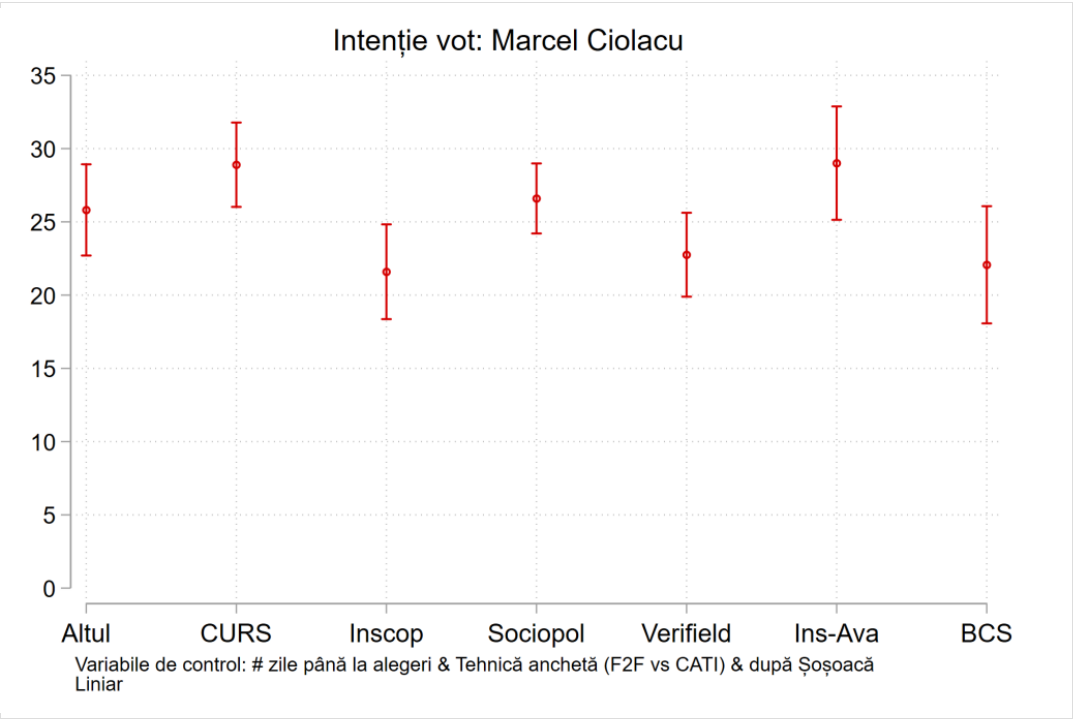
Folosind aceleași modele multivariate descrise în postările despre estimarea dinamicii intențiilor de vot, putem estima așa numitul efect de institut (house-effect). În condițiile prezente, o astfel de estimare e doar orientativă, pentru următoarele motive:

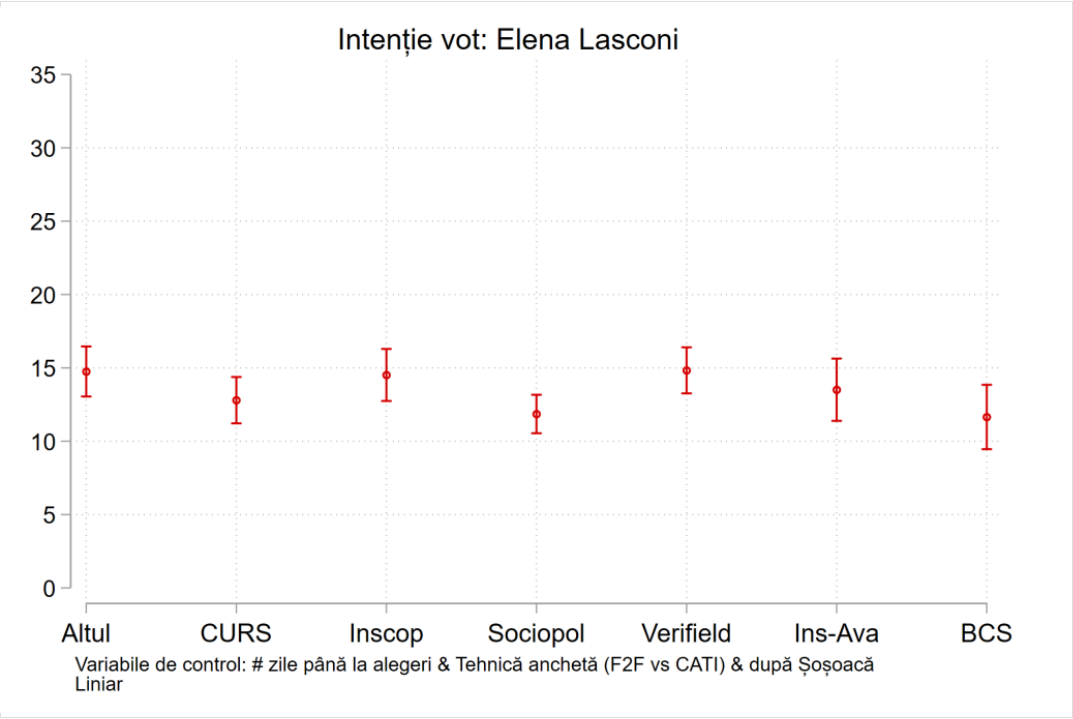
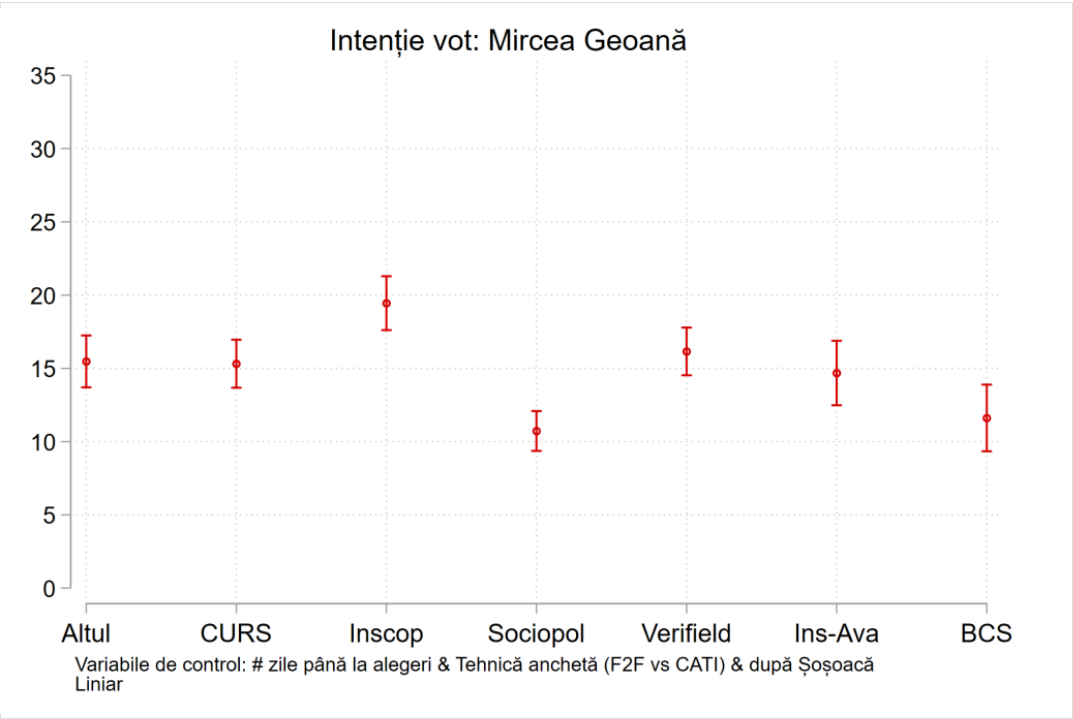
- numărul total de sondaje e redus,
- numărul de sondaje realizate în aproximativ aceeași perioadă e foarte mic,
- unele institute au făcut publice doar două sondaje,
- efectul de metodă se suprapune parțial cu efectul de institut și,
- **deocamdată, nu știm rezultatele finale, deci e posibil ca o deviere de la tendința medie să reflecte de fapt o estimare mai corectă.**

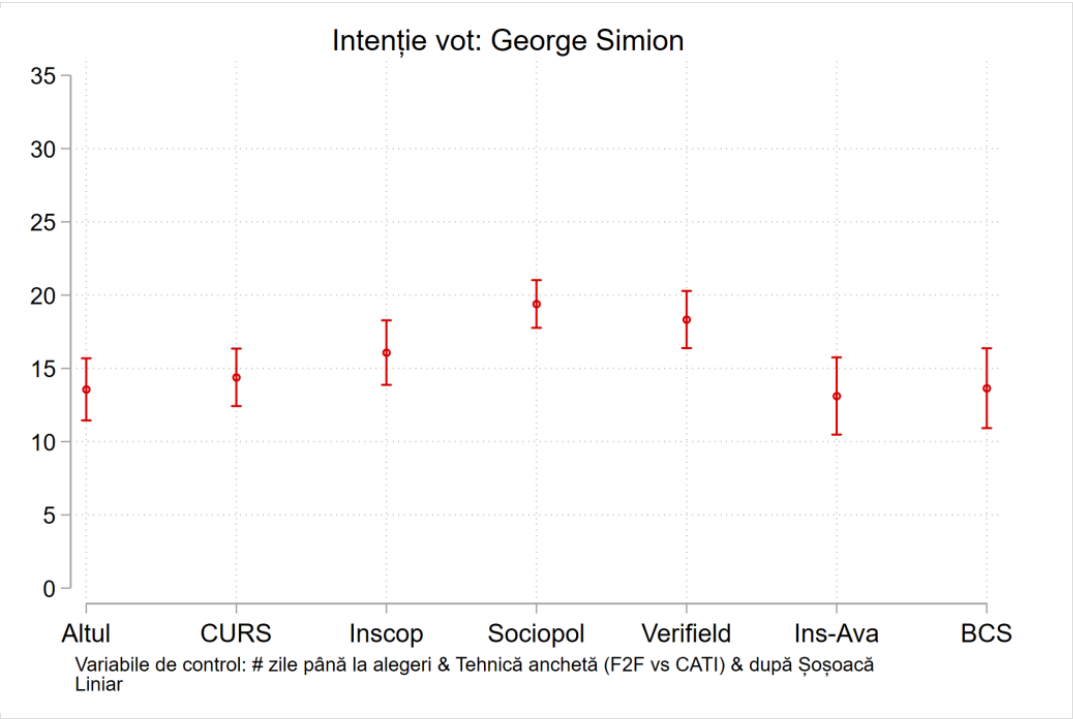
Fiecare grafic prezintă estimarea intenției de vot pentru un candidat, așa cum a fost ea prezisă de modelul regresie liniară multiplă, în funcție de institut, luând în calcul toate sondajele publicate în intervalul 1 iulie - 8 noiembrie 2024. Prin urmare, valorile din grafice reprezintă o estimare a intenției de vot considerând tot acest interval de timp, după eliminarea statistică a diferențelor induse de perioada de colectare a datelor, de tehnica de anchetă și de evenimente relevante (decizia CCR de a respinge

candidatura Dianei Șoșoacă). Punctele reprezintă estimarea, iar liniile intervalul de încredere asociat. Repet: **valorile din grafice nu reprezintă estimări ale intențiilor de vot de la un moment dat, nici estimări ale procentelor care vor fi obținute de candidați la alegeri; valorile reprezintă predicțiile unui model în care am controlat influența mai multor variabile (numărul de zile până la alegeri, tehnica de anchetă și impactul posibil al unor evenimente semnificative) cu scopul de a vedea dacă mai rămân diferențe între estimările făcute publice de către institutele de sondare.**

La ce e utilă o astfel de analiză? E normal să apară diferențe între estimările produse de diferite institute. Firesc, ne putem întreba cât de mari trebuie să fie aceste diferențe pentru a le considera puțin probabile din punct de vedere statistic. Diferențele dintre estimările diferitelor institute, așa cum sunt ele observate „cu ochiul liber,” pot avea ca sursă o serie de factori absolut legitimi (precum perioada de culegere a datelor, tehnica de anchetă, impactul unor evenimente). Prin modelare am eliminat impactul acestor factori. Dacă în urma controlului statistic tot mai apar diferențe semnificative statistic între estimările institutelor (adică liniile asociate institutelor nu se suprapun), aceste diferențe nu mai pot fi cauzate de factorii controlați, deci va trebui să căutăm alte explicații.







„House effects”: alegerile parlamentare 2024

Data: 2024-11-11

Cuvinte cheie: alegerile-parlamentare-2024, house-effects

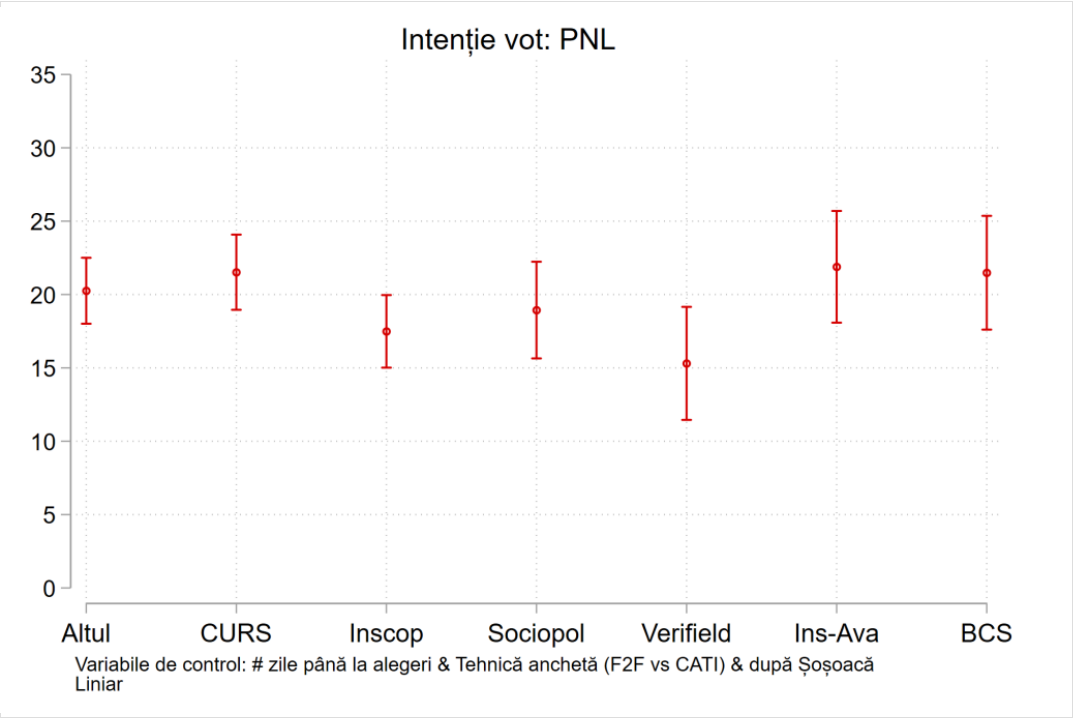
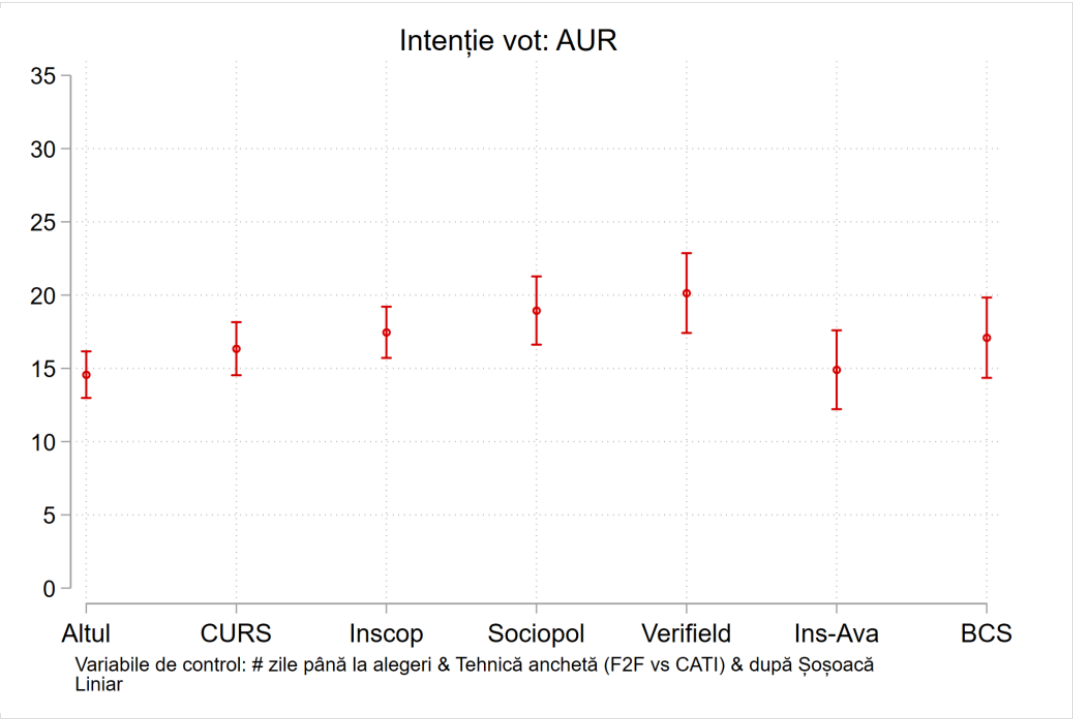
Folosind aceleași modele multivariate descrise în postările despre estimarea dinamicii intențiilor de vot, putem estima așa numitul efect de institut (house-effect). În condițiile prezente, o astfel de estimare e doar orientativă, pentru următoarele motive:

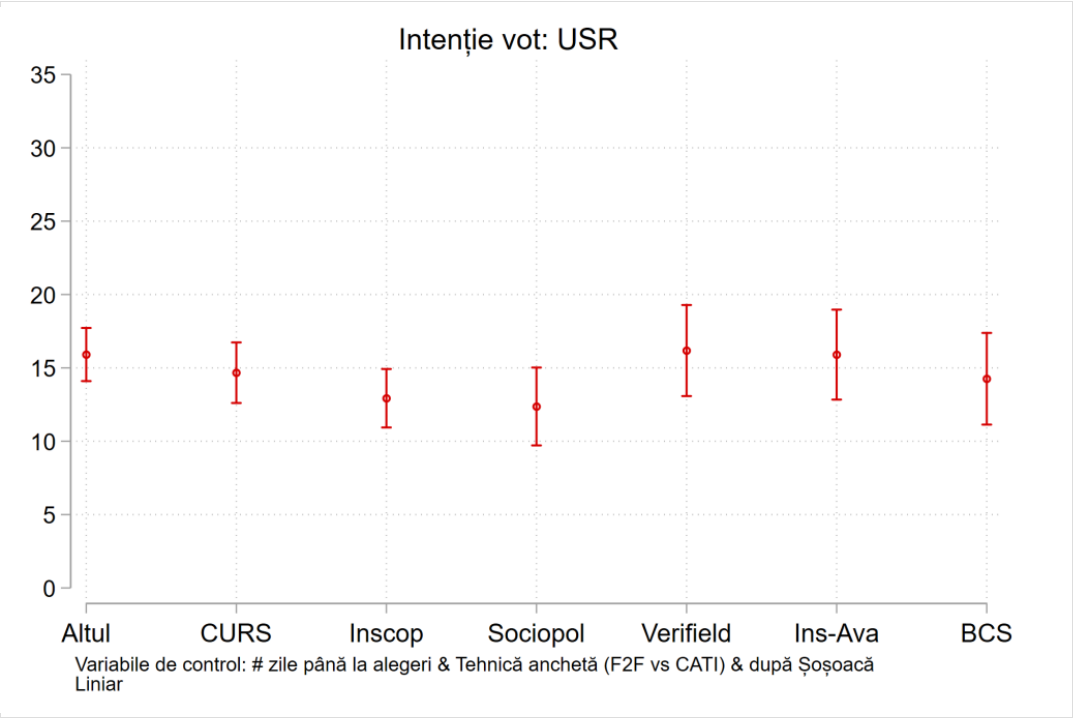
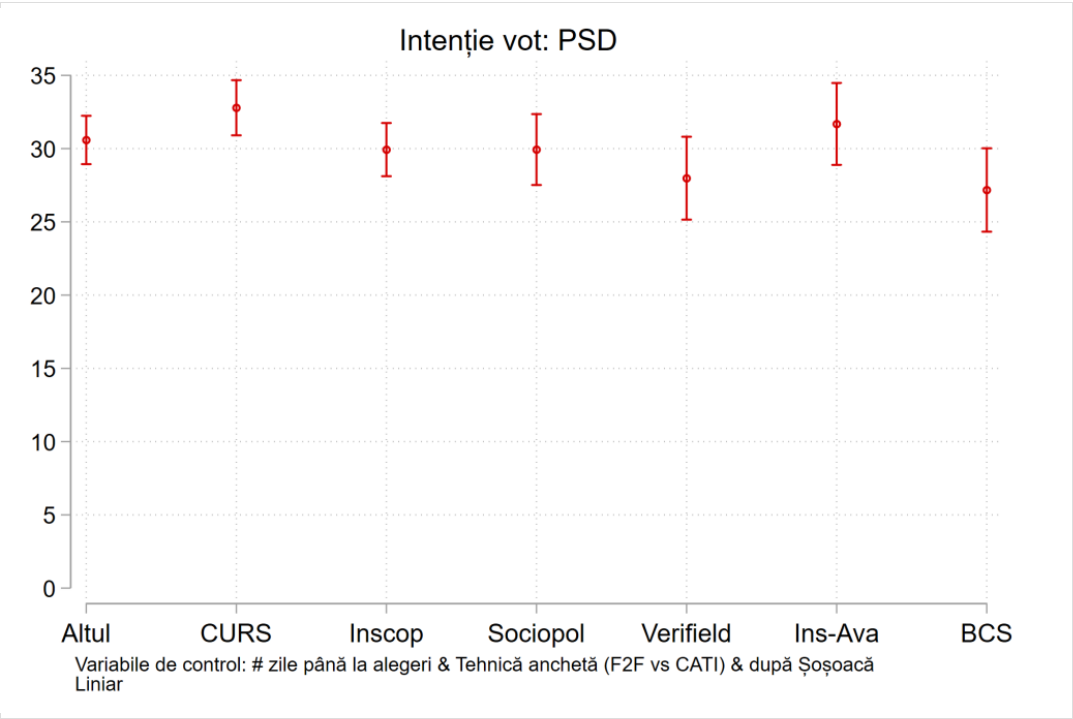
- numărul total de sondaje e redus,
- numărul de sondaje realizate în aproximativ aceeași perioadă e foarte mic,
- unele institute au făcut publice doar două sondaje,
- efectul de metodă se suprapune parțial cu efectul de institut și,
- **deocamdată, nu știm rezultatele finale, deci e posibil ca o deviere de la tendința medie să reflecte de fapt o estimare mai corectă.**

Fiecare grafic prezintă estimarea intenției de vot pentru un partid, așa cum a fost ea prezisă de modelul regresie liniară multiplă, în funcție de institut, luând în calcul toate sondajele publicate în intervalul 1 iulie - 8 noiembrie 2024. Prin urmare, valorile din grafice reprezintă o estimare a intenției de vot considerând tot acest interval de timp, după eliminarea statistică a diferențelor induse de perioada de colectare a datelor, de tehnica de anchetă și de evenimente relevante (decizia CCR de a respinge candidatura

Diane Șoșoacă). Punctele reprezintă estimarea, iar liniile intervalul de încredere asociat. Repet: **valorile din grafice nu reprezintă estimări ale intențiilor de vot de la un moment dat, nici estimări ale procentelor care vor fi obținute de partide la alegeri; valorile reprezintă predicțiile unui model în care am controlat influența mai multor variabile (numărul de zile până la alegeri, tehnica de anchetă și impactul posibil al unor evenimente semnificative) cu scopul de a vedea dacă mai rămân diferențe între estimările făcute publice de către institutele de sondare.**

La ce e utilă o astfel de analiză? E normal să apară diferențe între estimările produse de diferite institute. Firesc, ne putem întreba cât de mari trebuie să fie aceste diferențe pentru a le considera puțin probabile din punct de vedere statistic. Diferențele dintre estimările diferitelor institute, așa cum sunt ele observate „cu ochiul liber,” pot avea ca sursă o serie de factori absolut legitimi (precum perioada de culegere a datelor, tehnica de anchetă, impactul unor evenimente). Prin modelare am eliminat impactul acestor factori. Dacă în urma controlului statistic tot mai apar diferențe semnificative statistic între estimările institutelor (adică liniile asociate institutelor nu se suprapun), aceste diferențe nu mai pot fi cauzate de factorii controlați, deci va trebui să căutăm alte explicații.





Alegerile prezidențiale 2024 din Moldova: transferul voturilor

Data: 2024-11-13

Cuvinte cheie: alegerile-prezidentiale-2024-moldova

Chiar dacă în spațiul public din Republica Moldova mai apar uneori rapoarte cu privire la intențiile de vot sau alte analize primare, cel mai adesea datele primare de sondaj nu sunt disponibile. Din acest punct de vedere, alegerile prezidențiale din acest an nu sunt o excepție. Dimpotrivă. Și chiar dacă aș fi avut acces la micro-date, adică la baza de date cu răspunsurile alegătorilor la întrebările din sondaj, aș fi privit cu o oarecare rezervă răspunsurile respective (pentru motive ce țin de limite de natură metodologică și practică, respectiv de contextul socio-cultural nefavorabil unor investigații sociologice de acest fel).

În aceste condiții, o analiză a transferului de voturi pare sortită din start eșecului. Literatura de specialitate ne oferă însă o soluție: analiza/inferența ecologică. E vorba de o analiză realizată pe date la nivel agregat cu scopul de a produce inferențe cu privire la situația de la nivel individual. În domeniul studiilor electorale analiza este cunoscută sub numele de „voter transition matrix / probabilities” (poate fi realizată cu date la nivel agregat și/sau individual). Tot literatura de specialitate ne oferă și instrumentele necesare pentru a analiza date agregate (tehnica statistică și softul asociat). Chiar dacă o astfel de abordare are unele limite, în lipsa unei alternative mai bune, prezint aici două grafice care ne arată transferul de voturi de la turul 1 la turul 2 în cazul alegerilor prezidențiale din Moldova.

Pentru a reduce cât mai mult posibilitatea realizării unor inferențe incorecte, e important să folosim date situate la un nivel de agregare cât mai redus. Prin urmare, toate analizele sunt realizate pe datele oficiale la nivel de secție de vot (nivelul de agregare a datelor; datele cu privire la turul 2 au fost publicate ieri de CEC). Am considerat toate secțiile de vot, inclusiv cele din străinătate. Desigur, analiza se poate face și comparativ în funcție de zonă, raion, tip de localitate sau secție de vot. Intensitatea relațiilor dintre perechile de candidați este comunicată vizual prin intermediul intensității culorii și a mărimii fontului care indică valorile procentuale. O culoare mai intensă indică o pondere mai mare a voturilor transferate de la candidatul de la turul 1 la candidatul de la turul 2. Un font mai mare indică un număr absolut mai mare de voturi transferate. Cum estimările variază foarte puțin, mesajul transmis de cele două grafice este practic același:

- Finaliștii și-au păstrat cea mai mare parte din voturile primite la turul 1. Pare că, spre deosebire de Sandu, Stoianoglo și-a păstrat o pondere puțin mai mică din cota de la turul 1. Transferurile de voturi între cei doi sunt minime (1-3%) și foarte ușor în favoarea lui Sandu.
- Voturile lui Usatîi s-au împărțit aproximativ egal între cei doi finaliști (45% fiecare). Aproape 10% din votanții lui Usatîi nu au votat la turul 2.
- Stoianoglo a primit majoritatea voturilor lui Tarlev, Vlah, Furtună și Morari (și peste două treimi din voturile lui Chicu).
- Sandu a primit majoritatea voturilor lui Năstase, Ticu și Ulianovschi.
- Foarte important și decisiv pentru rezultatul alegerilor, Sandu a beneficiat de cea mai mare parte a voturile noi (~80%). Evident, cele mai multe au provenit din diaspora.
- Cel mai adesea, cei care nu au votat la turul 1 nu au votat nici la turul 2.

	t2_non_votanti	t2_bul_nev	t2_sandu	t2_stoianoglo	NET_EXITS	
t1_non_votanti	95.4	0.3	2.4	1.8	0.2	1477475
t1_bul_nev	0.9	28.8	68.9	1.1	0.3	18464
t1_sandu	1.0	0.2	97.6	1.0	0.1	656852
t1_stoianoglo	2.6	0.3	3.2	93.7	0.2	401215
t1_usatii	8.8	2.7	45.5	42.9	0.2	213169
t1_tarlev	0.8	0.2	0.8	98.0	0.2	49316
t1_vlah	0.2	0.1	0.0	99.6	0.2	83193
t1_chicu	1.7	2.4	25.8	70.0	0.2	31797
t1_nastase	0.1	0.1	99.6	0.0	0.2	9946
t1_ticu	0.2	0.1	99.5	0.0	0.2	14326
t1_furtuna	1.3	0.1	0.9	97.4	0.2	68778
t1_ulanovschi	0.3	5.5	91.5	2.4	0.1	7995
t1_morari	0.4	11.5	0.2	87.7	0.2	9444
NET_ENTRIES	4.8	0.2	77.0	18.1	0.0	118045
	1453118	20715	930139	750430	5613	

	t2_non_votanti	t2_bul_nev	t2_sandu	t2_stoianoglo	NET_EXITS	
t1_non_votanti	96.5	0.3	2.0	1.0	0.2	1477475
t1_bul_nev	0.3	30.0	69.0	0.4	0.3	18464
t1_sandu	0.4	0.1	98.6	0.8	0.1	656852
t1_stoianoglo	0.8	0.2	2.2	96.6	0.2	401215
t1_usatii	7.4	3.1	45.1	44.1	0.2	213169
t1_tarlev	0.2	0.1	0.4	99.1	0.2	49316
t1_vlah	0.1	0.0	0.0	99.7	0.2	83193
t1_chicu	0.6	2.6	25.4	71.2	0.2	31797
t1_nastase	0.0	0.0	99.7	0.0	0.2	9946
t1_ticu	0.1	0.0	99.7	0.0	0.2	14326
t1_furtuna	0.4	0.0	0.7	98.6	0.2	68778
t1_ulanovschi	0.1	5.8	93.1	0.8	0.1	7995
t1_morari	0.1	12.0	0.1	87.6	0.2	9444
NET_ENTRIES	4.3	0.3	80.3	15.1	0.0	118045
	1453118	20715	930139	750430	5613	

Dinamica intențiilor de vot: alegerile generale 2024

Data: 2024-11-16

Cuvinte cheie: alegerile-parlamentare-2024, alegerile-prezidentiale-2024, modelarea-intențiilor-de-vot

A trecut o săptămână de la cele două postări în care am prezentat primele estimări ale intențiilor de vot așa cum rezultă acestea din combinarea estimărilor sondajelor de opinie publice. Între timp au mai apărut în spațiul public câteva alte sondaje. În continuare, în cazul candidaților Ciucă și Simion (anterior era și Geoană pe această listă), apar diferențe mari, de până la 10 puncte procentuale, între estimările publicate de diferite institute de sondare a opiniei publice. Pare că fiecare institut a rămas fidel estimărilor realizate anterior. Foarte probabil, nu toate sondajele realizate sunt făcute publice. E posibil ca estimările produse de cele două categorii de sondaje, publice și private, să difere. De asemenea, e posibil ca sondajele care susțin o anumită distribuție a intențiilor de vot să fie totodată și sondajele care (nu) sunt făcute publice. Cum nu știu cum arată realitatea, voi ține cont de toate sondajele publicate, cu speranța că acestea, în medie, nu distorsionează intențiile de vot în (de)favoarea unui anumit competitor politic. După aflarea rezultatelor finale voi analiza măsura în care institutele de sondare au produs sau nu estimări distorsionate.

Dacă ținem cont doar de marja teoretică de eroare datorată eșantionării de tip simplu aleator (atenție, în practică selecția e de fapt tot mai puțin aleatoare), o diferență de 10 puncte procentuale este puțin probabilă chiar și atunci când comparația este realizată

corect, adică: (1) calculăm marja teoretică de eroare asociată fiecărei estimări prin raportare la numărul alegătorilor care au exprimat intenția de vot pentru un anumit candidat și nu la tot eșantionul și (2) calculăm suma celor două marje de eroare asociate estimărilor pe care le comparăm. Dacă diferența dintre cele două estimări este mai mare decât suma marjelor, putem spune că estimările sunt cel mai probabil diferite.

Pentru a înțelege mai bine cele descrise în paragraful anterior, propun următorul exemplu: Să presupunem că două institute de sondare, A și B, colectează fiecare, simplu aleator, date despre intenția de vot a 800 alegători (nu toți cei 1000 participanți la sondaj oferă un răspuns valid la întrebările de vot). Intenția de vot pentru candidatul X este 20% în sondajul A, respectiv 10% în sondajul B. Marjele teoretice de eroare sunt 2,8, respectiv 2,1 puncte procentuale. Cele două marje diferă pentru că valorile relativ la care sunt calculate diferă (dacă intenția de vot pentru partidul X ar fi fost 50%, marja de eroare ar fi crescut la 3,4 puncte procentuale). Prin urmare, estimarea obținută în sondajul A se află, cu o probabilitate de 95%, în intervalul 17,2;22,817,2; 22,8, iar estimarea obținută de sondajul B se situează în intervalul 7,9;12,17,9; 12,1. Altfel spus, dacă am realiza sondajul A de 100 de ori, în aceleași condiții, în 95 dintre sondaje am obține o estimare situată în intervalul 17,2;22,817,2; 22,8. Cum cele două intervale nu se suprapun, putem spune că intenția de vot pentru candidatul X estimată de sondajul A este diferită de intenția de vot estimată de sondajul B.

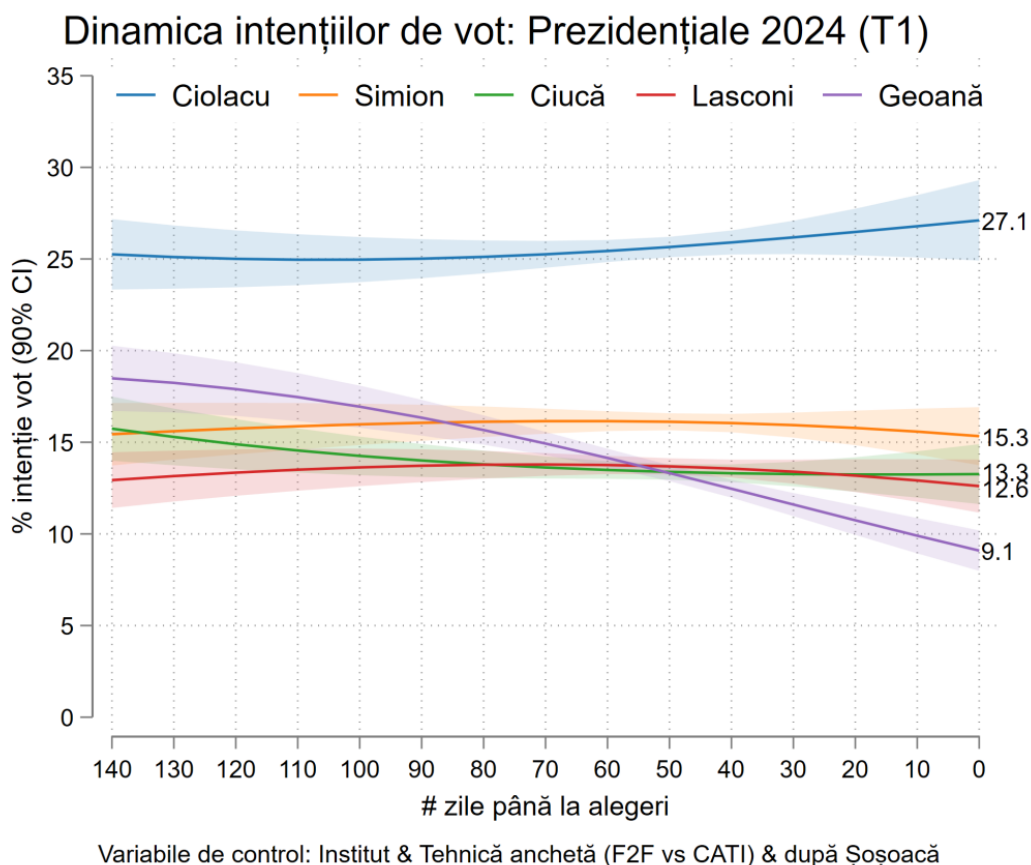
În practică, eroarea teoretică de eșantionare este cea mai mică problemă cu care se confruntă un institut de sondare. Pentru o serie de cauze pe care nu le enumăr aici, eșantioanele produse de majoritatea sondajelor sunt departe de a fi (strict) aleatoare. Simplu spus, cel mai adesea, eșantioanele rezultate nu reflectă fidel proporțiile de interes de la nivelul populației. Dacă luăm în calcul și faptul că nu toți cei selectați răspund la întrebarea de vot, rezultă că eroarea teoretică de eșantionare este doar o parte din eroarea totală asociată unui sondaj. **Pornind de la studii politice realizate în alte țări, putem estima că, în general, eroarea totală este aproximativ egală cu dublul erorii teoretice de eșantionare.** În lumina acestor informații, o diferență de 10 puncte procentuale

nu mai pare atât de neobișnuită. Totuși, atunci când o astfel de eroare apare sistematic, cu același semn, adică doar în (de)favoarea unui candidat, în cazul aceluiași institut de sondare, ar trebui să privim estimările respective cu rezervă. Suplimentar, specialiștii ar trebui să încerce să afle care e cauza acestor distorsiuni. Dacă diferențele observate nu sunt explicate de deciziile metodologice adoptate de institutul respectiv, se poate trece la următorii suspecti (publicarea doar a anumitor sondaje, finanțatorul / beneficiarul sondajului, conexiunile politice ale institutului de sondare etc.).

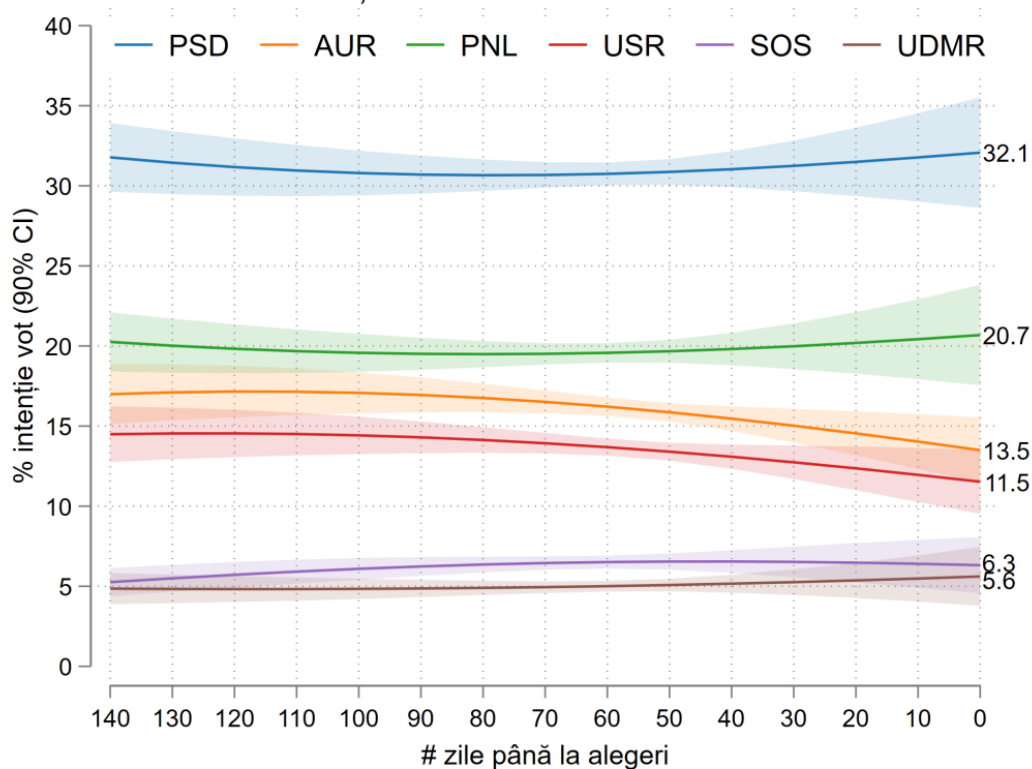
În cele ce urmează prezint o estimare a dinamicii intențiilor de vot la alegerile parlamentare, respectiv la turul 1 al alegerilor prezidențiale. Graficele au la bază estimările publicate în ultimele cinci luni de diferite institute de sondare a opiniei publice. Nu prezint rezultatele sondajelor ci doar estimările produse de modelul statistic implementat. Am inclus în model aceleași variabile de control: institutul care a realizat sondajul, tehnica de anchetă (CATI vs F2F) și eveniment 1 (colectarea datelor s-a realizat înainte sau după decizia CCR de a respinge candidatura Diane Șoșoacă). Spre deosebire de estimările anterioare, de această dată am ținut cont de toată perioada de colectare a datelor (nu doar de ultima zi), am estimat simultan intențiile de vot pentru toți competitorii mari, am estimat două modele ușor diferite dar am prezentat doar media estimărilor obținute, respectiv am extins estimarea până în ziua votului. Foarte important, aceste estimări ...

- țin cont doar de datele de sondaj,
- se referă doar la intențiile de vot de pe teritoriul României,
- nu iau în calcul calitatea estimărilor produse anterior de institute și nici eventualul bias al acestora,
- nu țin cont de posibilele efecte induse de (in)dezirabilitatea socială și/sau ratele diferite de participare la sondaje asociate electoratului diferiților candidați (astfel de factori pot influența toate estimările în același sens),

- Însurate dau mai puțin de 100%; diferența până la 100% e reprezentată de competitorii politici neincluși în model pentru motive evidente (nu au fost incluși în cele mai multe dintre sondajele publicate și/sau au apărut cu scoruri mici).



Dinamica intențiilor de vot: Parlamentare 2024



Variabile de control: Institut & Tehnică anchetă (F2F vs CATI) & după Șoșoacă

Zilele următoare intenționez să includ și alte sondaje, respectiv să realizez o estimare a intențiilor de vot care să țină cont, pe lângă datele de sondaj, și de (o parte dintre) limitele semnalate mai sus.

Alegerile generale 2024: „house effects”

Data: 2024-11-18

Cuvinte cheie: alegerile-parlamentare-2024, alegerile-prezidentiale-2024, house-effects

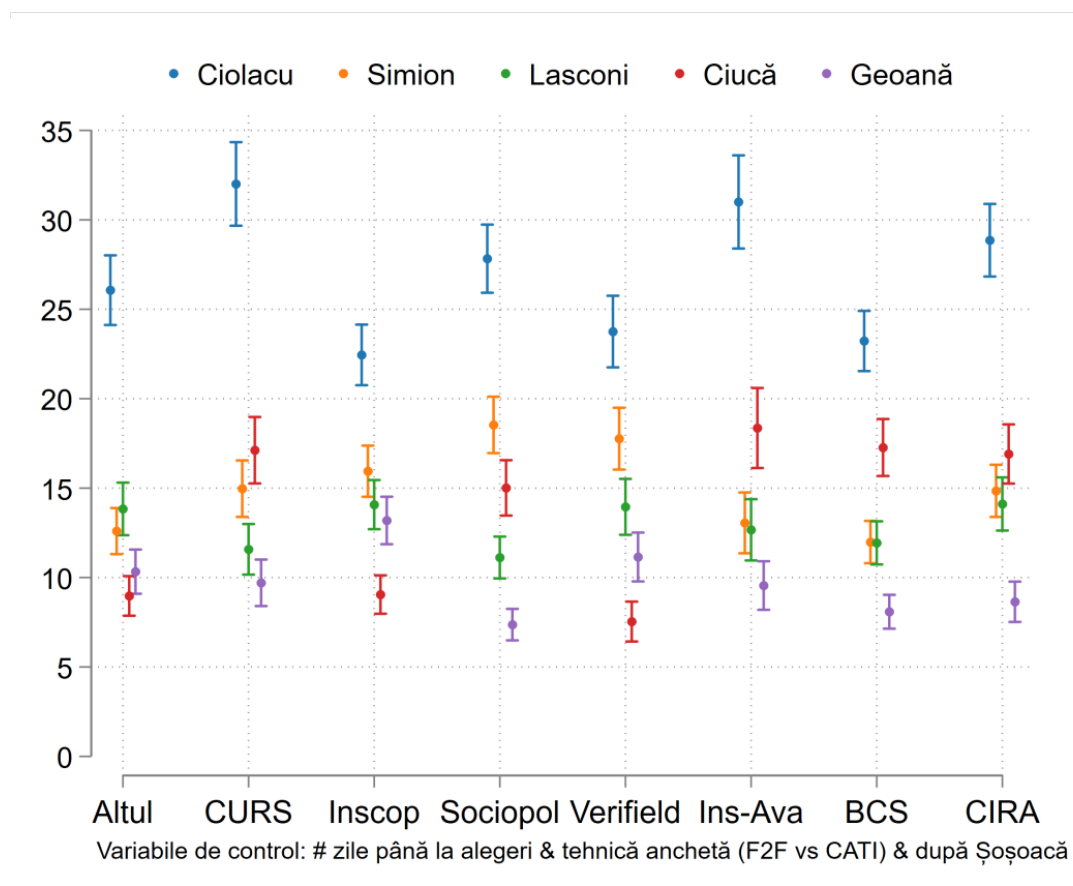
Probabil graficele din această postare pot fi înțelese mai bine dacă anterior au fost citite postările pe același subiect (vezi [aici](#) și [aici](#)). Graficele prezentate în această postare sunt produse în urma unei analize parțial diferite: am inclus sondajele apărute între timp, am ținut cont de toată perioada de colectare a datelor, nu doar de ultima zi de teren și am folosit o formulă ușor diferită de calculare a valorilor prezise. Valorile din grafice reprezintă valorile prezise de modelul statistic pentru ziua 20, adică 20 de zile înainte de data alegerilor. Desigur, pentru a putea estima cu adevărat „house-effects” e nevoie și de rezultatele oficiale.

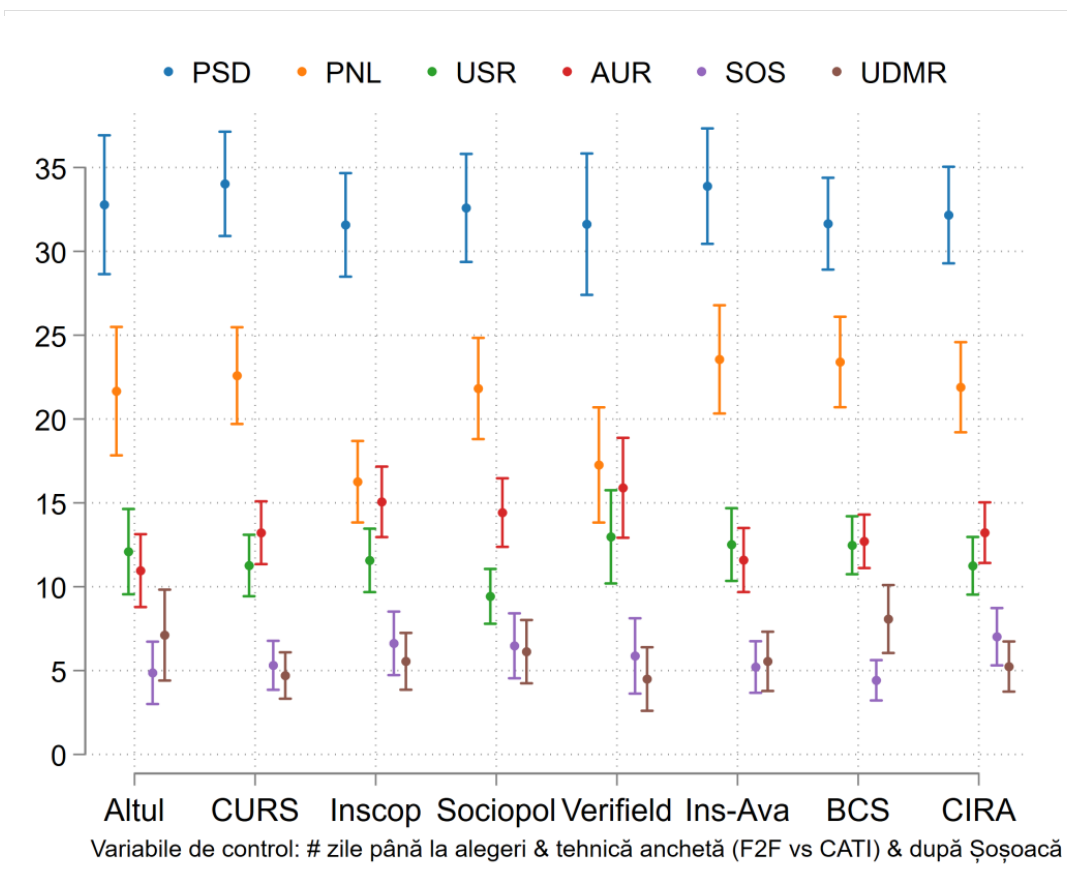
Atenție, nu am folosit aceeași culoare pentru a indica partidul și candidatul asociat acestuia (timpul e „di vină,” parțial lenea, respectiv probabilitatea mai mare de apariție a unor erori). Mult mai important: (1) în construcția modelului de predicție am ținut cont de toate sondajele publicate după 1 iunie, (2) distribuția în timp a sondajelor publicate diferă de la un institut la altul, (3) **valorile din grafice nu reprezintă nici intenții de vot așa cum apar în sondajele publicate, nici predicții ale rezultatelor alegerilor ci doar estimări ale intențiilor de vot la momentul $t-20$ (20 zile înainte de alegeri)**** în baza unui model multivariat cu scopul de a identifica eventualele diferențe dintre estimările produse de institute.** Simplu spus, am construit o serie de lumi alternative în care toate sondajele au fost colectate de același institut, cel

indicat la baza graficului, celelalte condiții considerate fiind aceleași, și am calculat cum arată „realitatea” privită prin lentila fiecărui institut în parte.

Cum graficele sunt suficient de clare, nu le mai „traduc” in extenso. Menționez doar două concluzii:

- diferențele dintre estimările asociate candidaților la alegerile prezidențiale sunt mai mari comparativ cu cele asociate partidelor;
- în cazul unor perechi, sentimentele de „iubire” și „ură” apar destul de clar; va trebui să așteptăm rezultatul alegerilor pentru a vedea cu adevărat care institute au reușit să ignore „sentimentele” și care nu; desigur, e posibil să aflăm că realitatea este pe alocuri colorată, pe alocuri gri.





Alegerile generale 2024: F2F vs. CATI

Data: 2024-11-19

Cuvinte cheie: alegerile-parlamentare-2024, alegerile-prezidentiale-2024, f2f-vs-cati

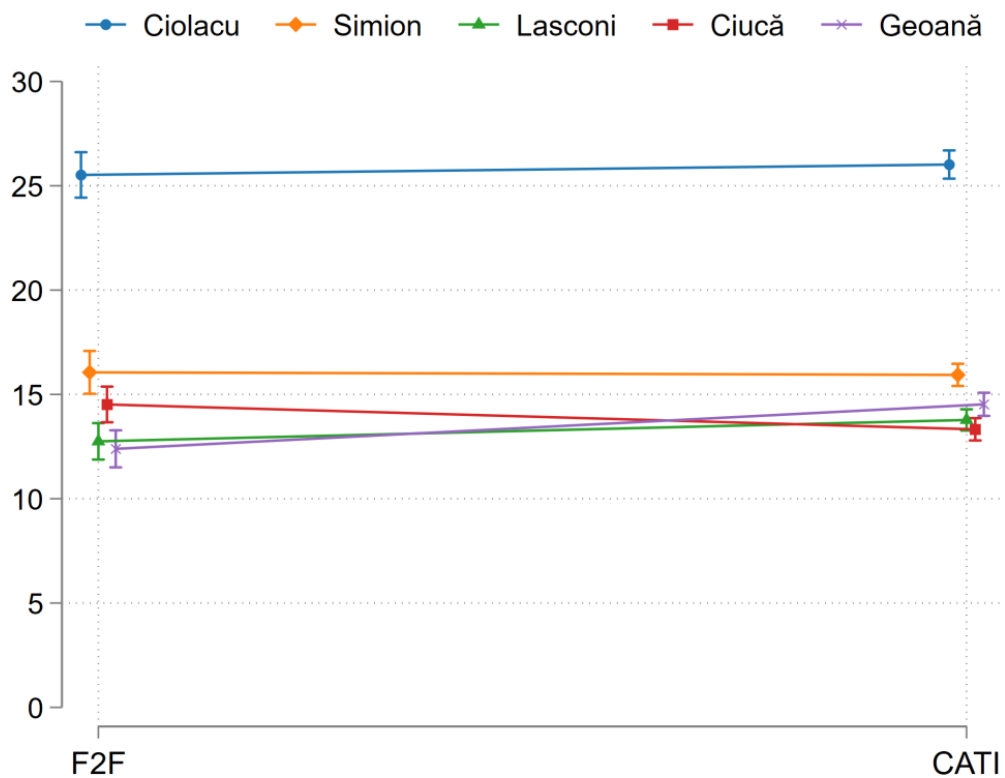
Tema acestei postări este aparent pur metodologică. Însă, comparația în funcție de tehnica de anchetă (Față-în-Față vs. CATI) este utilă în contextul estimării intențiilor de vot, respectiv a diferențelor dintre estimările publicate de institute. Logica de estimare este descrisă într-o [postare anterioară](#), deci nu o mai reiau. Faptul că „toată lumea observă cu ochiul liber” diferențe între estimările produse de cele două tehnici de anchetă (F2F și CATI) nu înseamnă neapărat că aceste diferențe chiar există. E foarte posibil ca diferențele să aibă de fapt o altă cauză, de exemplu institutul care colectează datele (deciziile metodologice luate, respectiv implementarea lor practică în principal). Desigur, pot fi și alți factori. Pentru a vedea care este sursa mai probabilă a acestor diferențe apelăm la analiza multivariată. În traducere: „Atunci când ținem cont de o serie de factori (institutul care a realizat sondajul, numărul de zile până la alegeri, diferite evenimente importante), diferențele în funcție de tehnica de anchetă apar în continuare?”

În cazul acestei analize, cel mai adesea nu au apărut diferențe semnificative statistic în funcție de tehnica de anchetă: nicio diferență în cazul partidelor, respectiv trei diferențe în cazul candidaților (Ciolacu, Lasconi și Geoană). Foarte important, **niciuna dintre diferențe nu a fost mai mare de două puncte procentuale**. Prin urmare, din punct de vedere practic putem

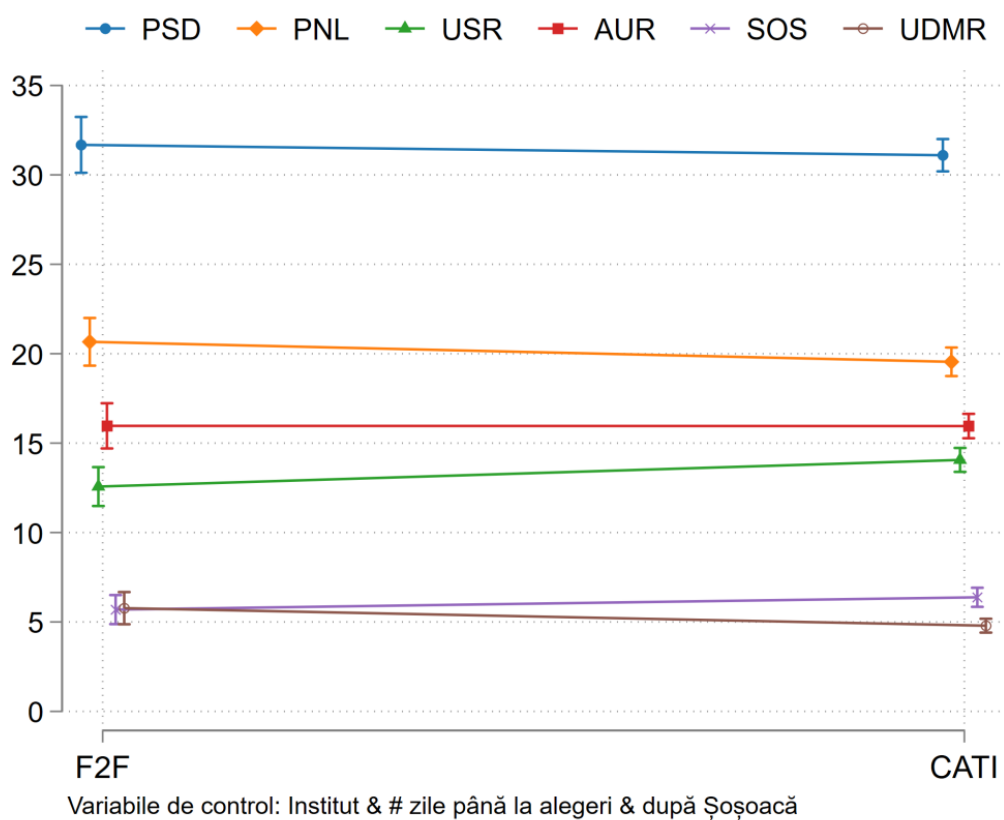
spune că mai degrabă nu apar diferențe între estimările produse de cele două tehnici de anchetă.

- în cazul sondajelor de tip CATI apar scoruri puțin mai mari pentru Lasconi (+1), Geoană (+2), USR (+1) și SOS (+1);
- în cazul sondajelor de tip F2F apar scoruri puțin mai mari pentru Ciucă (+1), PNL (+1) și UDMR (+1).

Câteva limite ale analizei: numărul de sondaje realizate față-în-față (F2F) este foarte mic; sondajele F2F sunt realizate cel mai adesea de aceleași institute (unele institute nu au făcut niciun sondaj F2F); în cazul anumitor candidați (ceva mai puțin pentru partide), estimările variază mult în funcție de institutul care a realizat sondajul; în aceste condiții tehnica statistică nu poate identifica suficient de clar care e sursa diferențelor (dacă tehnica de sondaj chiar produce estimări diferite); pentru aceasta ar fi fost de preferat o abordare de tip experimental.



Variabile de control: Institut & # zile până la alegeri & după Șoșoacă



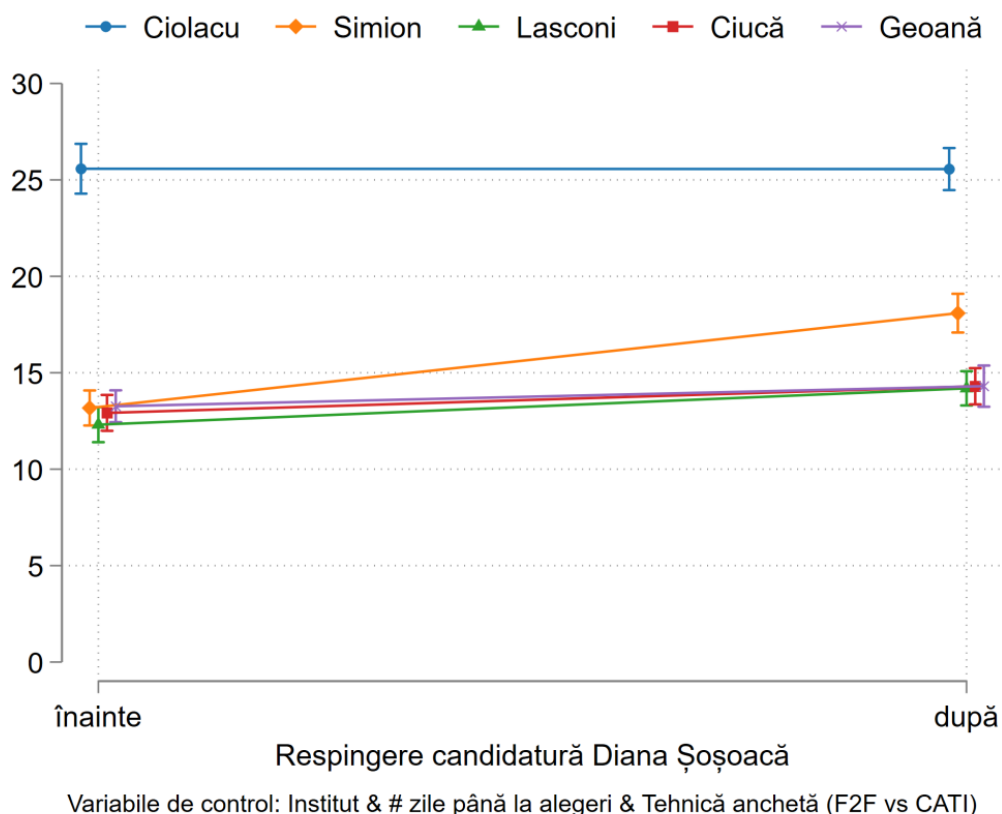
Prezidențiale 2024: intențiile de vot înainte și după decizia CCR

Data: 2024-11-19

Cuvinte cheie: alegerile-prezidentiale-2024, decizia-ccr, dinamica-intentiilor-de-vot, rdd

Eliminarea din competiția electorală a candidatei Diana Șoșoacă a produs destul de multă vâlvă în media și, prin comparație, probabil mai puțină la nivelul alegătorilor. Pornind de la datele de sondaj disponibile, foarte limitate, încerc totuși să estimez care a fost impactul deciziei CCR. Mai exact, încerc să identific candidații care, probabil, au beneficiat / pierdut puncte procentuale în urma deciziei CCR. Pentru a răspunde la această întrebare propun două strategii de analiză.

Prima strategie este descrisă într-o [postare anterioară](#), deci nu o mai reiau. Conform estimărilor obținute prin această metodă, principalul beneficiar a fost George Simion. Intenția de vot pentru acesta a crescut cu aproximativ 5 puncte procentuale. Fiind vorba despre un candidat din opoziție, era oarecum de așteptat să asistăm la o creștere a intențiilor de vot pe măsură ce se apropie ziua alegerilor. În cazul celorlalți candidați din opoziție se observă o creștere de cel mult două puncte procentuale. Prin urmare, doar în cazul lui Simion creșterea e sensibil mai mare (5 puncte). Făcând diferența, rezultă că, în urma deciziei CCR, Simion a câștigat aproximativ 3 puncte procentuale.



A doua strategie apelează la ceea ce în literatura de specialitate se numește Regression Discontinuity Design (RDD). RDD urmărește să evalueze ipoteza prezenței unui punct de ruptură într-o serie statistică. În acest caz e vorba despre numărul de zile rămase până la data alegerilor, iar punctul de interes este ziua în care a fost luată decizia CCR. Folosind această abordare, putem analiza dinamica intenției de vot asociate unui candidat în funcție de un eveniment de referință. Mai exact, putem vedea astfel dacă, pentru a descrie matematic această dinamică, putem ignora evenimentul de interes și e suficientă o singură funcție sau, dimpotrivă, nu putem ignora acel eveniment și avem nevoie de două funcții. Dacă observăm că apar creșteri sau scăderi ale intențiilor de vot în proximitatea evenimentului analizat, putem identifica, cu anumite limite, candidații care au pierdut / câștigat puncte procentuale ca urmare a deciziei CCR.

Atenție! Și această abordare e departe de o logică strict experimentală. Mai mult, datele disponibile sunt limitate: nu au fost publicate sondaje care să acopere zilele din proximitatea datei de 5 octombrie, data hotărârii CCR, iar estimările variază uneori mult în funcție de institutul care a realizat sondajele.

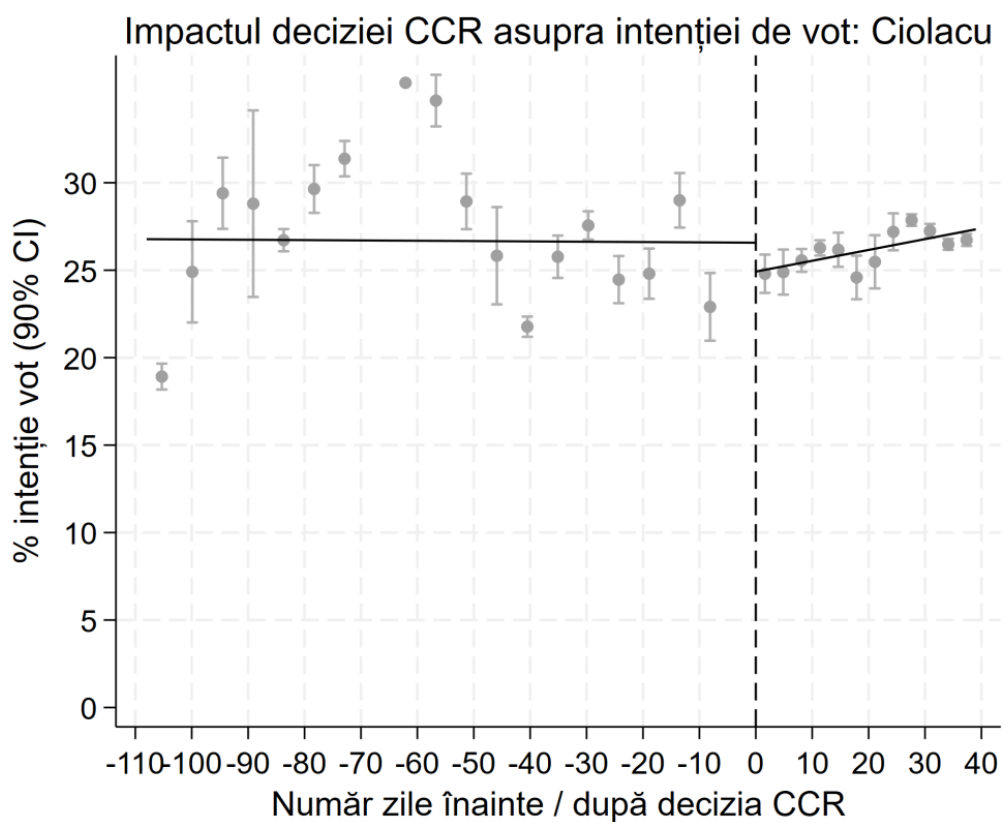
Graficele de mai jos prezintă diferențele observate în urma analizei RDD. În partea din stânga a fiecărui grafic este prezentată situația din perioada anterioare deciziei CCR, iar în dreapta situația de după. Punctele gri și intervalele de încredere asociate reprezintă distribuția estimărilor intențiilor de vot conform sondajelor, iar linia neagră indică tendința medie a distribuției respective (dată fiind variabilitatea estimărilor din sondaje am preferat să sintetizez datele folosind un model simplu, liniar).

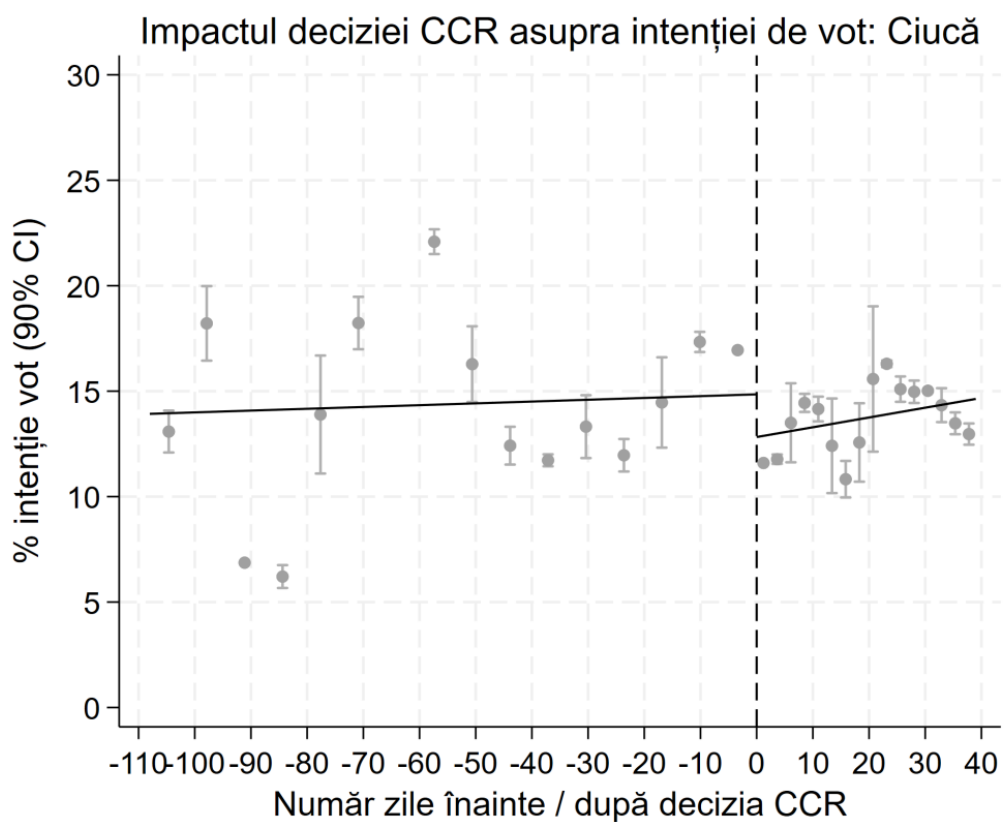
Pornind de la aceste rezultate, pot fi extrase câteva concluzii cu privire la fiecare candidat major:

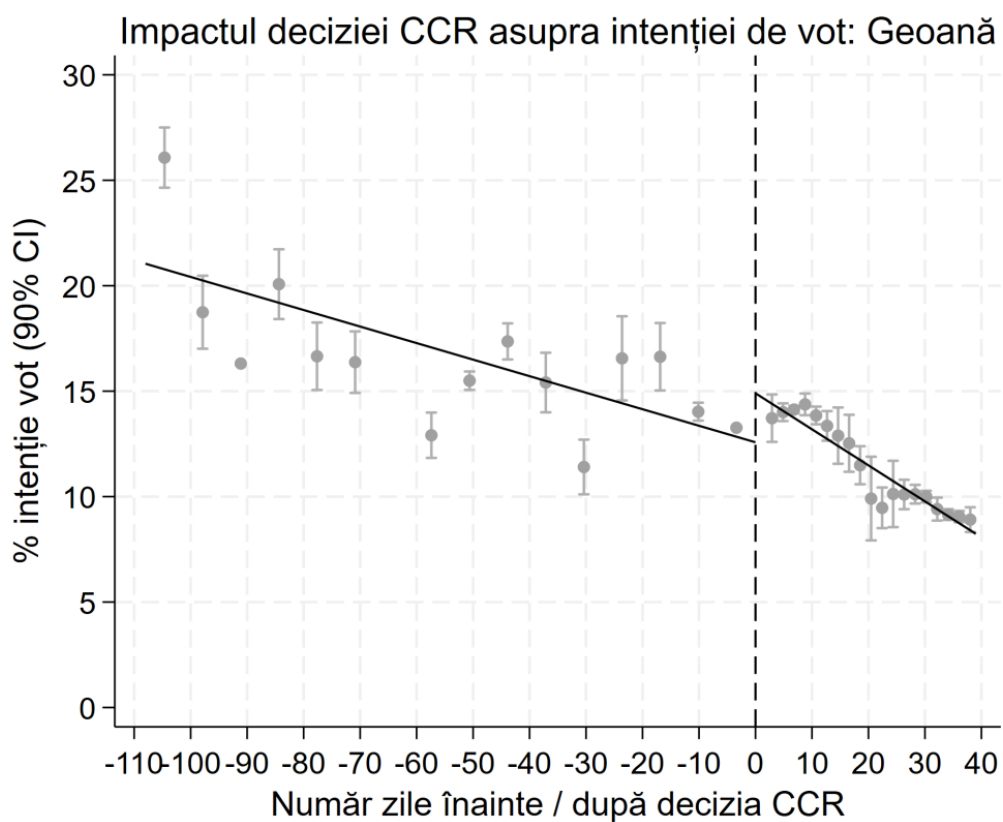
- **Ciolacu:** în prima perioadă observăm o variabilitate foarte mare a estimărilor și lipsa unui trend; cauzele probabile ale variabilității estimărilor sunt numărul mare de zile până la alegeri și indecizia cu privire la propria candidatură; în a doua perioadă observăm o variabilitate foarte mică a estimărilor (institutele „se pun de acord” cu privire la intenția de vot asociată acestui candidat) și un trend ușor ascendent; cel mai probabil, decizia CCR nu a avut un impact asupra intențiilor de vot pentru acest candidat;
- **Ciucă:** în prima perioadă observăm o variabilitate foarte mare a estimărilor și un trend ușor ascendent; în a doua perioadă variabilitatea estimărilor scade, dar e în continuare mare (estimările sunt dependente de institutul care a realizat sondajul); trendul este ușor ascendent; cel mai probabil, decizia CCR nu a avut un impact asupra intențiilor de vot pentru acest candidat;
- **Geoană:** în prima perioadă observăm o variabilitate foarte mare a estimărilor (institutele tind să producă estimări diferite) și un trend descendent; în a doua perioadă variabilitatea estimărilor este mică (institutele

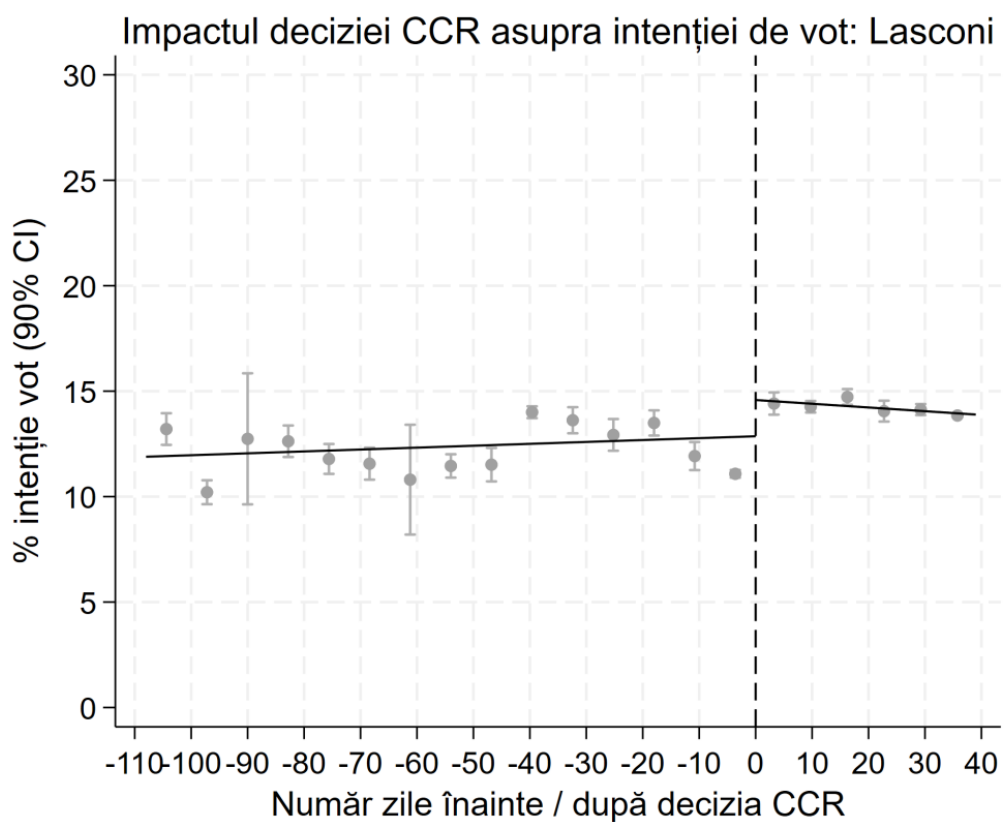
tind să producă estimări similare), iar trendul negativ se accentuează; cel mai probabil, decizia CCR nu a avut un impact asupra intențiilor de vot pentru acest candidat;

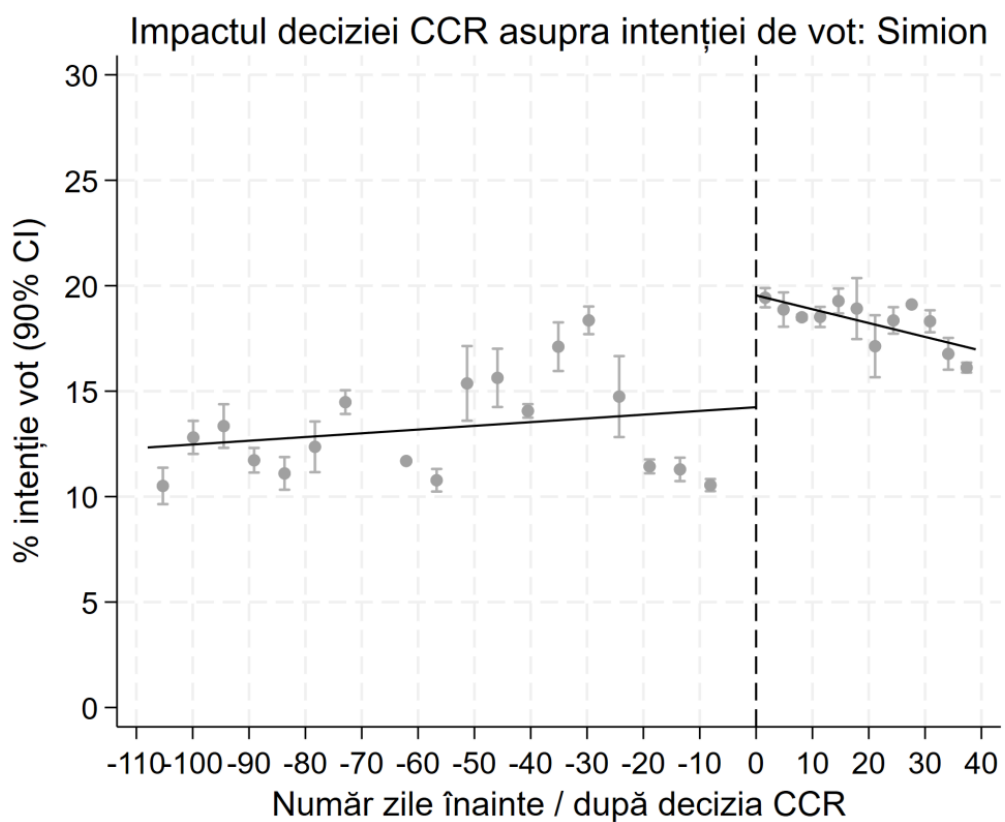
- **Lasconi:** în prima perioadă observăm o variabilitate redusă a estimărilor și un trend ușor ascendent; în a doua perioadă variabilitatea estimărilor scade și mai mult, iar trendul tinde să fie ușor descendent; cel mai probabil, decizia CCR nu a avut un impact asupra intențiilor de vot pentru acest candidat;
- **Simion:** în prima perioadă observăm o variabilitate în creștere a estimărilor (pe măsură ce ne apropiem de ziua alegerilor institutele produc estimări tot mai diferite, ceea ce e destul de ciudat) și un trend ascendent; în a doua perioadă variabilitatea estimărilor scade mult, iar trendul devine negativ; cel mai probabil, decizia CCR a determinat o creștere de aproximativ cinci puncte procentuale a intențiilor de vot pentru Simion; ulterior, pe măsură ce alți candidați similari (Georgescu, Terheș) au devenit mai vizibili, o parte a votanților Diane Șoșoacă s-au reorientat spre aceștia; totuși, **aproximativ trei puncte procentuale au rămas în continuare la Simion, însă e posibil ca această „zestre” să mai scadă până la alegeri;**
- **Alt candidat:** în prima perioadă observăm o variabilitate foarte mare a estimărilor și un trend ascendent; în a doua perioadă variabilitatea estimărilor scade, dar rămâne totuși mare; trendul ascendent se accentuează puțin; firește, în perioada imediat următoare deciziei CCR intenția de vot pentru această categorie de candidați a scăzut, dar apoi situația a revenit destul de repede la starea anterioară deciziei CCR.

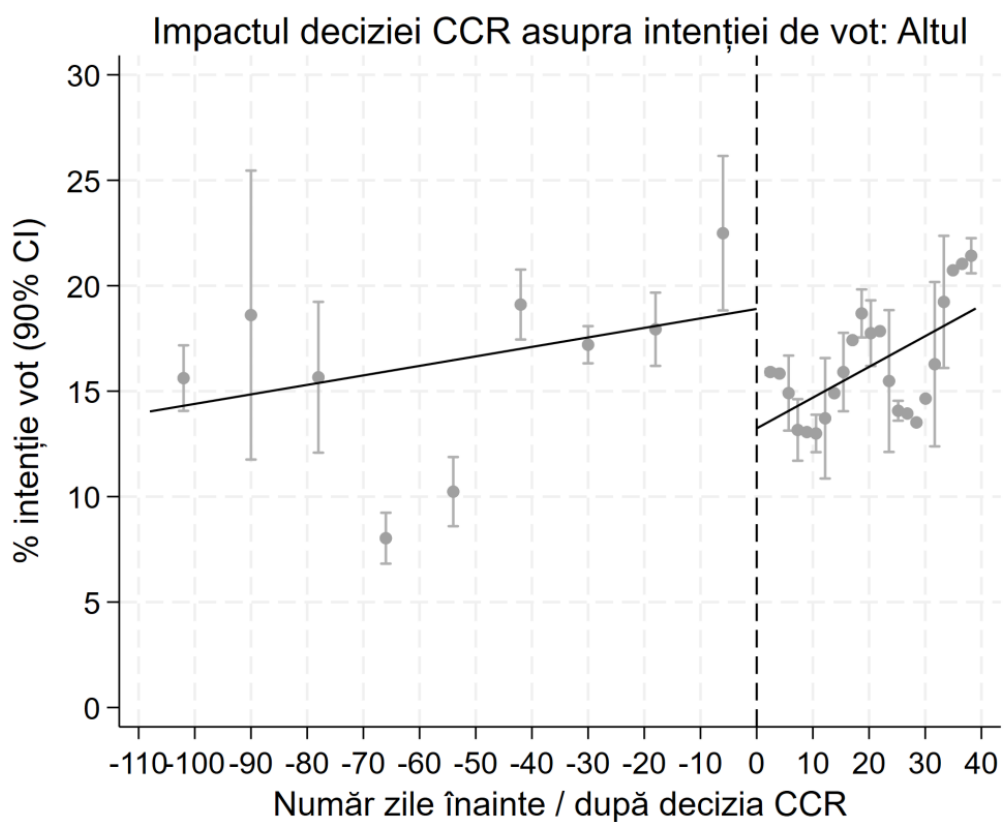












Lupta din adâncurile sondajelor

Data: 2024-11-20

Cuvinte cheie: alegerile-prezidentiale-2024, house-effects

Anul acesta au fost publicate o serie de sondaje pe tema intențiilor de vot la alegerile prezidențiale. În anumite cazuri, diferențele dintre estimările produse au fost destul de mari. În mod normal, pe măsură ce ne apropiem de ziua alegerilor, diferențele ar trebui să scadă pentru că:

- lista candidaților e (aproape) finală;
- majoritatea alegătorilor știu cine sunt candidații (cei majori cel puțin);
- tot mai mulți alegători se decid dacă și cu cine votează; desigur, aproximativ 15% din votanți se decid în ultima săptămână, probabil procentul fiind mai mare în cazul acestei competiții;
- volumul de informații și gradul de expertiză specifică acestor alegeri cresc;
- numărul de sondaje și numărul de subiecți cresc și ele.

Pentru a afla dacă estimările tind să se stabilizeze, am realizat, evident, un grafic. Pe scurt, am procedat așa:

- am calculat pentru fiecare candidat o medie a estimărilor produse de sondaje într-o anumită perioadă de timp (medie mobilă 10);

- similar, am calculat diferența medie asociată valorilor obținute la pasul anterior (abaterea standard a mediilor mobile);
- am împărțit a doua valoare la prima (coeficientul de variație);
- am obținut astfel, pentru fiecare candidat, la diferite momente de timp, o măsură a gradului în care estimările variază; foarte important, această măsură nu depinde de mărimea absolută a procentelor asociate unui candidat, prin urmare valorile obținute pot fi comparate între ele;
- pentru fiecare candidat am indicat tendința probabilă a valorilor obținute (mărimea coeficienților de variație) folosind o funcție polinomială de grad 2.

Dacă lucrurile se întâmplă așa cum ar trebui să se întâmple, adică diferențele dintre estimările institutelor să scadă în timp în cazul tuturor candidaților, **ar trebui ca tendințele (liniile) din grafic să fie descendente și să convergă spre zero. Concret, ar trebui ca liniile să fie orientate spre partea din dreapta jos a graficului.**

În realitate acest lucru se întâmplă doar în cazul candidaților Ciolacu și Lasconi. Aceștia sunt și candidații în cazul cărora apare cel mai mare consens cu privire la estimări. În cazul lui Ciolacu s-a plecat de la un consens scăzut, care a crescut rapid. În cazul candidatei Lasconi, inițial consensul a fost mai degrabă mediu și a crescut ușor în timp. Important e că există un consens ridicat în ambele cazuri. Tendințele asociate celorlalți trei candidați sunt „originale,” fiecare într-un mod diferit:

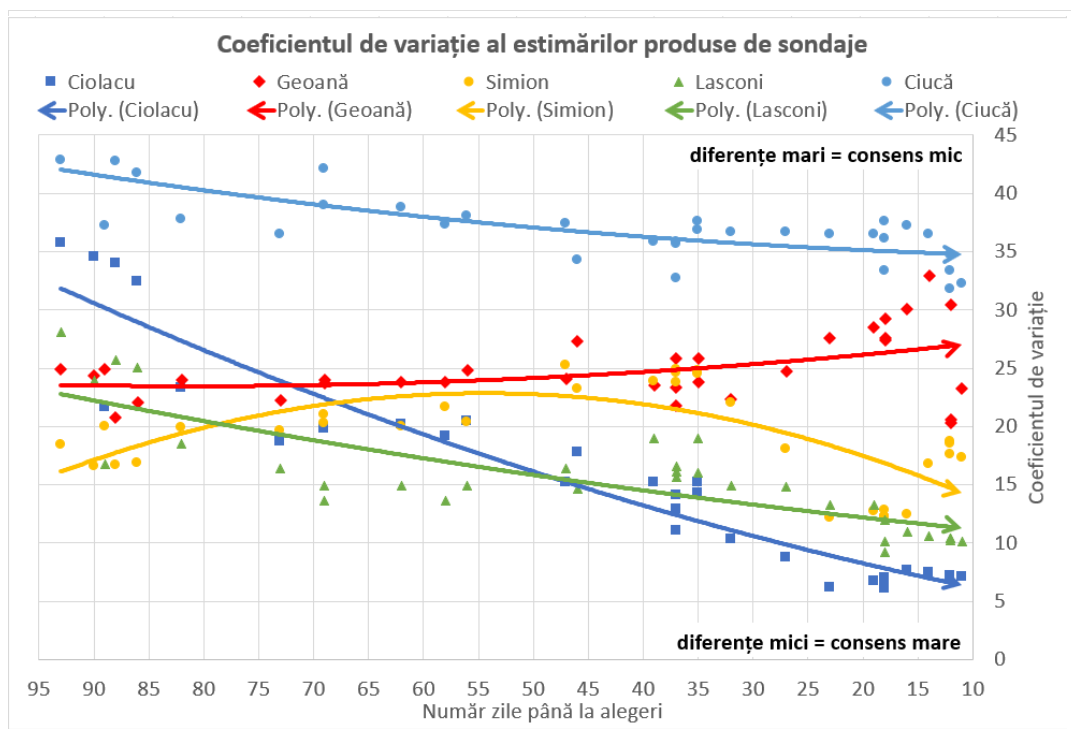
- **Simion:** Se pleacă de la un consens mai degrabă ridicat al estimărilor (diferențele dintre estimările realizate de institute sunt mai degrabă mici; la început cel mai mare consens apărea în cazul lui Simion), apoi consensul scade, după care revine la starea inițială.
- **Ciucă:** Se pleacă de la un consens mic (institutele publică estimări foarte diferite) care deși crește ușor, rămâne în continuare redus. Ciucă este candidatul în cazul căruia

se observă cel mai scăzut consens pe tot parcursul perioadei analizate.

- **Geoană:** Pentru toată perioada de referință se observă un consens mediu, relativ stabil, care, surprinzător, pare să scadă ușor spre final. Într-un clasament al consensului scăzut, Geoană se situează pe poziția secundă, după Ciucă.

Într-o traducere simplă, **institutele încă „deliberează” cu privire la scorul pe care îl vor obține candidații Ciucă și Geoană.** E dificil de spus care sunt cauzele care au dus la această situație. Ipotezele posibile pe care le văd sunt următoarele:

- Dificultatea estimării intențiilor de vot este mai mare în cazul acestor candidați. O astfel de situație e posibilă dacă electoratele celor doi sunt mai puțin omogene comparativ cu restul candidaților.
- Metodologii diferite. Dacă a fost așa, de ce efectele au apărut doar în cazul celor doi? Au fost luate decizii care au favorizat un candidat? Poate întrebarea de vot a fost pusă diferit sau într-un context diferit. Poate ponderarea datelor a fost făcută diferit. Poate datele publicate au fost produse în urma unei modelări, iar modelele au fost diferite.
- Finanțatorul. Ar trebui să ținem cont de cine a plătit pentru fiecare sondaj, respectiv de conexiunile dintre finanțatori și candidați. Deși nu ar trebui să conteze, uneori poate contează.
- „Publication bias.” Au fost publicate doar sondajele favorabile? E posibil ca anumite clauze incluse în contractele încheiate între finanțatori și institute să ofere exclusiv finanțatorilor libertatea de a decide dacă și ce rezultate sunt publicate.



Alegerile prezidențiale 2024, turul 1: o estimare a intențiilor de vot

Data: 2024-11-22

Cuvinte cheie: alegerile-prezidentiale-2024, modelarea-intențiilor-de-vot

În trei postări anterioare ([aici](#), [aici](#) și [aici](#)) am prezentat o serie de analize care mă ajută acum să produc o estimare a intențiilor de vot. În cele trei analize am folosit date din sondajele publice: intențiile de vot exprimate ca procente, perioada de colectare a datelor, tipul anchetei și institutul care a realizat sondajul. Eu doar am modelat statistic datele disponibile.

În această analiză reiau problema estimării intențiilor de vot folosind același tip de modelare, dar adaug la modelul inițial câteva variabile. Scopul acestui demers este de a testa dacă modelul produce o predicție cât mai apropiată de rezultatul alegerilor de duminică. Când spun rezultatul alegerilor mă refer la **rezultatul final, cel care se referă la toate voturile exprimate (voturile din țară și din diaspora)**. Astfel, am inclus în noul model informații despre ...

- o serie de **factori care pot influența calitatea estimărilor din sondaje**; mă refer aici la fenomene precum dezirabilitatea socială, spirala tăcerii, segmente electorale asociate potențial unui candidat și care au rate reduse de participare la sondaje; desigur, astfel de efecte nu apar în cazul tuturor candidaților și nici nu prea pot fi estimate corect fără date experimentale; totuși, pornind de la o serie de situații anterioare

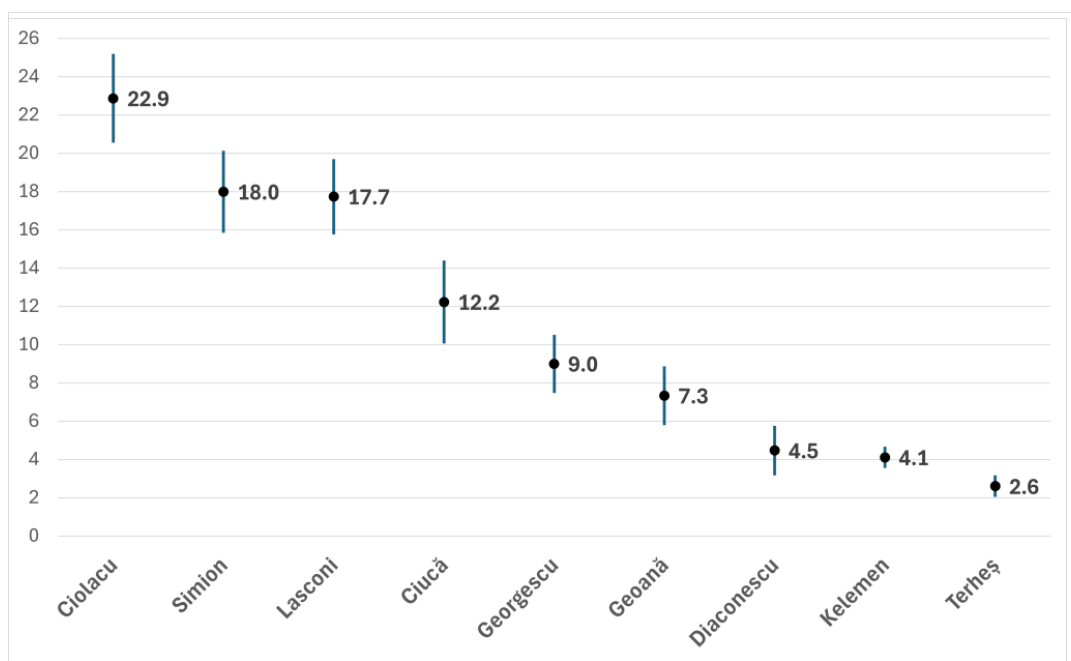
similare, în cazul candidaților majori, am modificat foarte puțin estimările produse prin modelarea datelor din sondaje; modificările au fost de cel mult o jumătate de punct procentual, în sus sau în jos, funcție de direcția în care acționează efectele considerate;

- **diferite evenimente recente** care ar putea influența preferințele de vot; aici am ținut cont de anunțurile făcute de Orban și Băsescu, ambele în favoarea candidatei Lasconi; și în acest caz modificările au fost mai mici de o jumătate de punct;
- **votanții din diaspora**; chiar dacă în cazul acestora nu avem date de sondaj, intențiile de vot pot fi approximate pornind de la distribuția observată a voturilor în cazul unor competiții electorale recente;
- **numărul votanților**; am luat în calcul prezența estimată de diferite sondaje în combinație cu numărul votanților în cazul unora dintre competițiile electorale anterioare; pe baza acestor informații, ținând cont și de percepțiile publice cu privire la candidați, am anticipat o prezență ceva mai redusă comparativ cu alegerile prezidențiale din 2019.

Toate informațiile / estimările / aproximările cu care lucrez au un anumit grad de incertitudine. Prin urmare, nu am folosit valori fixe ci distribuții posibile asociate acestor valori. Am considerat că distribuțiile au o formă de clopot (sunt normale). Procedând astfel, modelarea ține cont de incertitudinile asociate valorilor de input. În cazul distribuțiilor asociate candidaților majori am folosit grade diferite de incertitudine, acestea fiind estimate pe baza variațiilor din sondaje. Un model care nu ține cont de aceste aspecte riscă să producă estimări care (1) sunt prea sigure în relație cu informația disponibilă sau, dimpotrivă, (2) sunt prea imprecise, deci prea puțin informative sau (3) asociază același grad de incertitudine a estimării tuturor candidaților, ceea ce evident nu e cazul. Am produs astfel o mulțime de „lumi alternative.” Apoi, pentru a vedea care sunt scenariile cele mai probabile și/sau dacă putem diferenția între diferite scenarii, am pus informația rezultată într-o formă grafică.

Departe de mine gândul că model propus aici va produce estimări perfecte. Mă aștept însă ca scorul obținut de fiecare candidat să fie undeva în interiorul intervalului de încredere (distribuției estimate). Cu cât e mai aproape de centrul acesteia, cu atât modelarea e mai bună. Nu e cazul să vorbesc acum despre limitele modelului și nici despre posibilele soluții. Voi ține cont de toate acestea în viitor. Dacă aș fi avut acces la micro-date (bazele de date) sau dacă informațiile cu privire la sondaje ar fi fost complete sau dacă institutele ar fi fost mai puțin dependente de actorii politici, modelul statistic ar fi arătat altfel, iar estimările ar fi fost probabil (parțial) diferite. De asemenea, **ar fi fost util ca toate institutele să publice estimările finale**, cele făcute pe baza datelor colectate în ultima săptămână. Din păcate, cel puțin două institute lipsesc. Pentru că, în cazul unor candidați, apar diferențe mari între estimările institutelor, includerea în analiză doar a unora dintre sondajele realizate în ultima săptămână poate distorsiona estimările finale.

Pentru a nu deruta cititorii mai puțin obișnuiți cu grafice de tip distribuție de densitate, prezint inițial rezultatul modelării într-o formă grafică simplă. În acest grafic, cercurile pline indică estimările intențiilor de vot, iar liniile indică incertitudinea estimărilor, adică intervalul în care ne așteptăm, cu o probabilitate de 95%, să se situeze rezultatul final.



Pe scurt, mesajul acestui grafic e următorul:

- Ciolacu e aproape sigur pe primul loc, dar scorul obținut va fi probabil mai mic decât cel așteptat (asta și ca urmare a numărului mic de voturi pe care mă aștept să-l primească din diaspora).
- Pe baza datelor existente nu pot spune cine va ocupa locul secund. Poate fi oricare dintre candidații Simion sau Lasconi. Totuși, sunt șanse puțin mai mari să fie Simion. Balanța va fi înclinată într-o parte sau alta în funcție de distribuția deciziilor alegătorilor care s-au decis sau se vor decide zilele acestea, duminică inclusiv. Cum numărul lor nu e deloc mic (~ 20%), oricare dintre cei doi candidați poate câștiga locul doi. Și nu în ultimul rând, ordinea finală va depinde de cum vor vota în realitate votanții din diaspora (pentru că nu am avut date de sondaj, în acest caz am folosit aproximări cu un grad relativ mai mare de incertitudine).

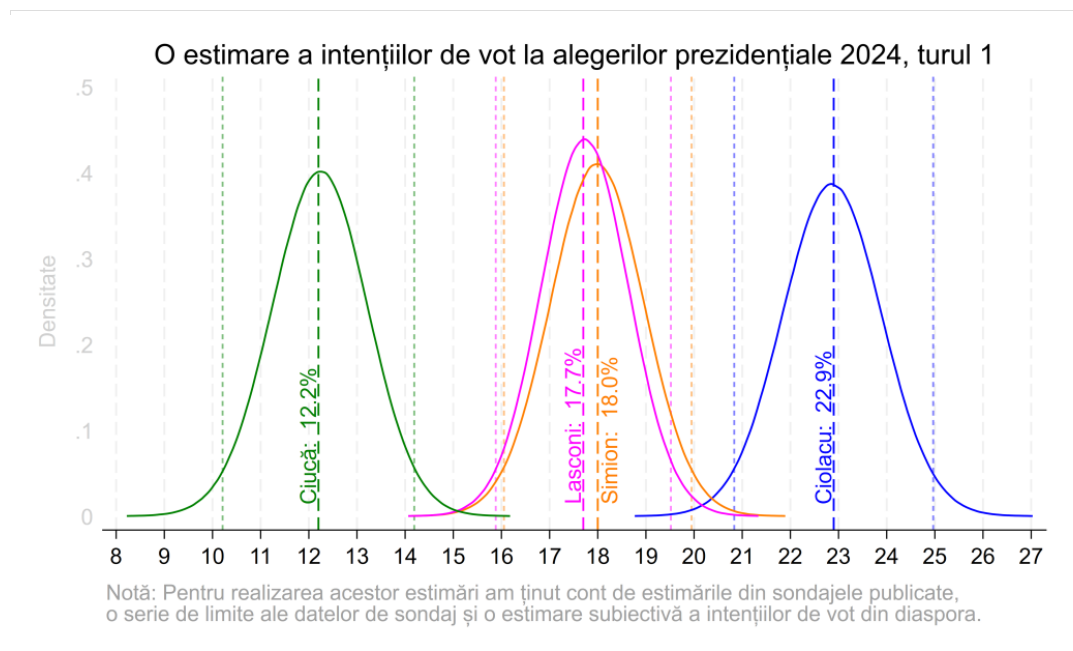
- Ciucă ocupă locul patru. Lucrurile se mai pot schimba, dar probabilitatea să piardă această poziție în favoarea următorului clasat e totuși mică.
- Pe locurile următoare se situează, în ordine descrescătoare, Georgescu, Geoană, Diaconescu, Kelemen și Terheș. Diferența dintre Georgescu și Geoană este în marja de eroare, deci practic ei ar putea fi la egalitate (la fel în cazul lui Diaconescu și Kelemen). Pentru că Birchall a apărut cu un scor mai ridicat într-un singur sondaj (AtlasIntel), nu am inclus-o în acest clasament.

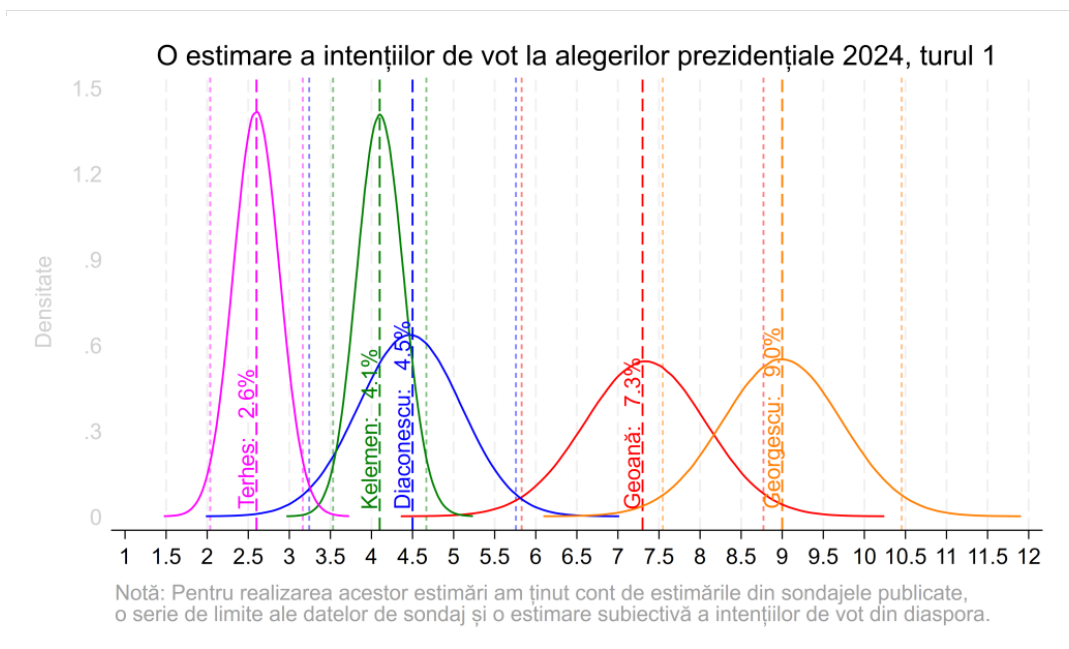
Pentru cei interesați să afle mai multe informații, mai tehnice, prezint rezultatele și sub forma densităților de probabilitate asociate estimărilor. Dat fiind faptul că numărul candidaților e destul de mare, am organizat estimările în două grafice. Un grafic de tip densități de probabilitate ne indică în principal următoarele:

1. Care este procentul „cel mai probabil” din voturi care va fi obținut de un candidat. Răspunsul e indicat de valoarea de pe axa orizontală care corespunde punctului de inflexiune al curbei. Am indicat valoarea respectivă și printr-o linie punctată însoțită de menționarea numelui candidatului și a procentului estimat. Această valoare nu are nicio legătură cu cât de înaltă e curba. De exemplu, chiar dacă Terheș are aproape cea mai înaltă curbă, are intenția cea mai mică de vot.
2. Cât de sigură este estimarea procentului „cel mai probabil” amintit mai sus: cu cât curba este mai înaltă (sau mai puțin lată), cu atât estimarea este mai precisă. De exemplu, estimarea pentru candidatul Terheș este mai precisă comparativ cu estimarea pentru candidatul Geoană.

3. Care este intervalul în care se situează, cu o probabilitate de 95%, estimarea respectivă. Răspunsul e indicat doar vizual prin cele două linii întrerupte care însoțesc fiecare estimare. Acestea au aceeași culoare, mai puțin intensă. De exemplu, candidatul Terheș va obține, cel mai probabil, undeva între 2.0 și 3.2% din totalul voturilor.

Dacă modelul propus aproximează realitatea suficient de bine, ar trebui ca procentul obținut de fiecare candidat să se regăsească undeva în zona definită de cele două linii verticale marcate printr-o culoare mai puțin intensă. Cu cât scorul obținut de fiecare candidat va fi mai aproape de linia verticală asociată acelui candidat, cu atât modelul va fi mai bun.





Atenționare: cele două grafice au scale diferite în cazul ambelor axe, prin urmare putem compara vizual doar curbele și ariile din același grafic.

Poate vă întrebați de ce înălțimea curbelor diferă atât de mult, respectiv de ce unele sunt mai extinse pe orizontală. Tocmai pentru că incertitudinea asociată estimărilor diferă de la un candidat la altul. În cazul unor candidați, Ciucă de exemplu, există o variabilitate mai mare a valorilor de input, adică a estimărilor din sondaje. Prin urmare, e firesc ca această variabilitate (incertitudine) să fie reflectată de rezultatul modelării. Similar, pentru că nu am avut date, aproximările votului din diaspora sunt și mai incerte (totuși, pentru că sunt relativ mai puțini votanți din diaspora, impactul asupra estimării finale e mic).

Și la final, pe scurt, o idee creată. Parcă prea au fost liniștite alegerile acestea. Oare cum ar fi fost dacă unul dintre candidații cu o zestre ceva mai mare de voturi s-ar fi retras din competiție în ultimul moment. Evident, nu putea fi vorba de Ciolacu, Lasconi sau Ciucă. Toți sunt susținuți de partide care au nevoie de voturile suplimentare pe care un candidat calificat în finală le-ar aduce la parlamentare. Puțin probabil ca Diaconescu sau Georgescu să se fi

retras (Diaconescu a declarat de mai multe ori că nu va face acest lucru, iar Georgescu pare să fie în „misiune”; nu se știe dacă pentru acum sau pentru ciclul electoral viitor). Mai rămânea Geoană. Dacă inițial păreau să existe tatonări între PSD și Geoană, ulterior mariajul nu a mai avut loc. S-au poate că totuși a fost cândva o nuntă, undeva departe de ochii publicului. De ce? Pentru un motiv foarte simplu: un candidat independent care să „fure” câteva puncte procentuale de la candidații Ciucă și Lasconi e mană cerească pentru PSD și Ciolacu. Două, trei puncte nu sunt multe, dar sunt suficiente pentru a-i ține departe de locul doi pe Ciucă și Lasconi. Mai rămân PNL și USR. Ambele partide ar fi avut de câștigat, separat și/sau împreună, de pe urma voturilor lui Geoană (doar o parte a acestora; totuși electoratul unui candidat nu poate fi mutat dintr-un coș în altul de pe azi pe mâine). Mă miră faptul că nu au ajuns la o înțelegere reciproc avantajoasă. Ce s-ar fi întâmplat dacă Geoană s-ar fi retras, cu sau fără asumare publică, și și-ar fi îndemnat electoratul să-l voteze pe unul sau pe oricare dintre candidații Lasconi sau Ciucă? În urma unei astfel de decizii, lupta pentru locul doi ar fi fost probabil tot strânsă, dar Lasconi ar fi fost în fața lui Simion. Iar dacă Lasconi ajunge în finală, sunt șanse să revedem, pentru a patra oară, același film electoral. Pentru că retragerea lui Geoană nu a avut loc, iar o retragere azi e foarte puțin probabilă, ne întoarcem la scenariul unu: „Și dacă a avut loc totuși o nuntă!”

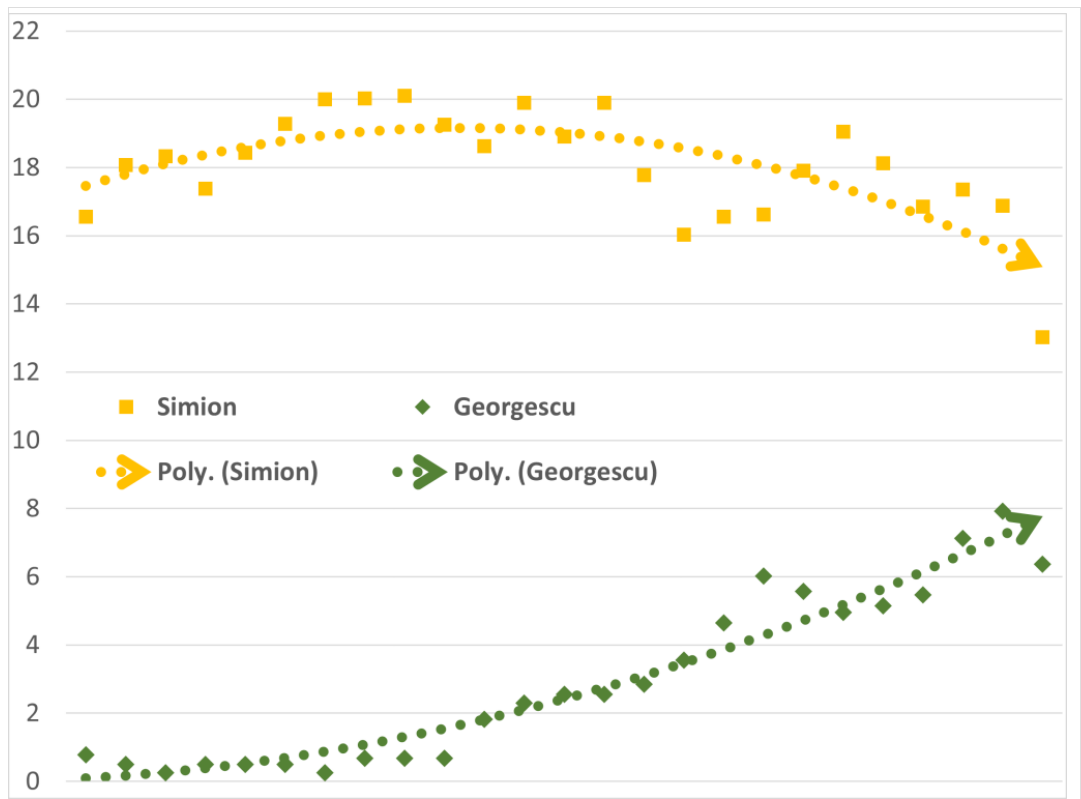
30

Călin file de poveste

Data: 2024-11-23

Cuvinte cheie: alegerile-prezidentiale-2024, dinamica-intentiilor-de-vot

Se apropie seara, prin urmare vă voi spune o poveste scurtă, fără intenția de a vă adormi. Dimpotrivă.



Sfârșit!

Pentru cei fără somn, continui totuși cu un scurt supliment.

Dacă datele de sondaj prezentate în ultima zi de campanie sunt corecte, povestea aceasta poate fi cheia alegerilor. Și nu doar a celor din turul 1.

Strict pe baza datelor agregate (medii mobile), că altele nu am, pare că există un segment de electorat aflat în căutarea unui anumit tip de mesaj și a unei anumite figuri politice. În absența unei etichete mai potrivite, dacă doriți, îl putem numi suveranist (deși cred că termenul simplifică prea mult agregarea de valori și credințe care caracterizează acest electorat). E probabil ca acesta să fie electoratul Diane Șoșoacă care, după decizia CCR, s-a trezit fără reprezentare. La început, o parte din acești votanți s-au mutat în barca lui George Simion (vezi [aici](#)), parțial și Cristian Terheș, apoi, după aproape o lună, cu sau fără ajutor, s-au trezit pe bune și s-au orientat spre Călin Georgescu. Cu siguranță, nu au fost singurii care au fost mișcați de noua figură politică. Destul de probabil, succesul poveștii a fost potențat și de decizia BEC prin care i se impunea candidatului Călin Georgescu să elimine din mediul online toate postărilor electorale.

Povestea continuă, doar că, de această dată, **avem doi competitori care aleargă pe culoare care se suprapun parțial, iar unul din ei, deși a pierdut puține puncte, se află în situația să piardă probabil mai mult!**

Update estimări T1

Data: 2024-11-23

Cuvinte cheie: alegerile-prezidentiale-2024, modelarea-intentiilor-de-vot

Ieri a fost ceva buluceală la sondaje, ceea ce e foarte bine. Însă, nu e prea bine că, în unele cazuri, apar diferențe cam mari între estimări. Nu e prea bine nici că nu au fost publicate ultimele sondaje realizate de institutele care au trecut prin focul mai multor competiții electorale naționale (Inscop, Sociopol, Avangarde). Datele lor ar fi fost utile pentru a contrabalansa eventualele probleme asociate sondajelor realizate de celelalte institute (AtlasIntel și Verifield). În cazul acestora din urmă nu există sondaje realizate cu ocazia unor competiții naționale anterioare, deci nu putem estima nivelul de încredere în datele produse.

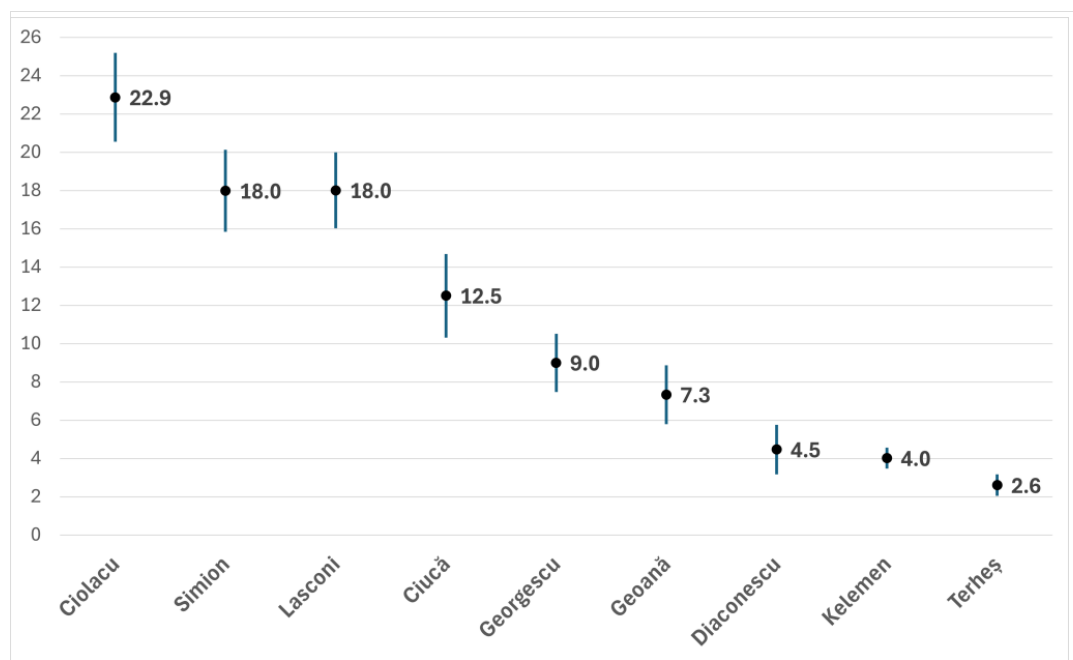
Refac aici simularea de ieri incluzând și ultimul sondaj (AtlasIntel), chiar dacă am unele rezerve cu privire la eșantioanele produse prin metoda lor (MRP nu poate corecta chiar totul). Nu spun că metoda nu funcționează foarte bine în SUA sau în București (vezi [aici](#)). Însă, am motive să suspectez că, deocamdată, funcționează mai puțin bine în cazul alegerilor generale din România (sau din alte contexte similare). La Europarlamentare au avut erori mari (vezi [aici](#)). Da, e drept, au făcut sondajul cu multă vreme înainte de alegeri, deci e foarte posibil ca situația să fi fost alta la acel moment.

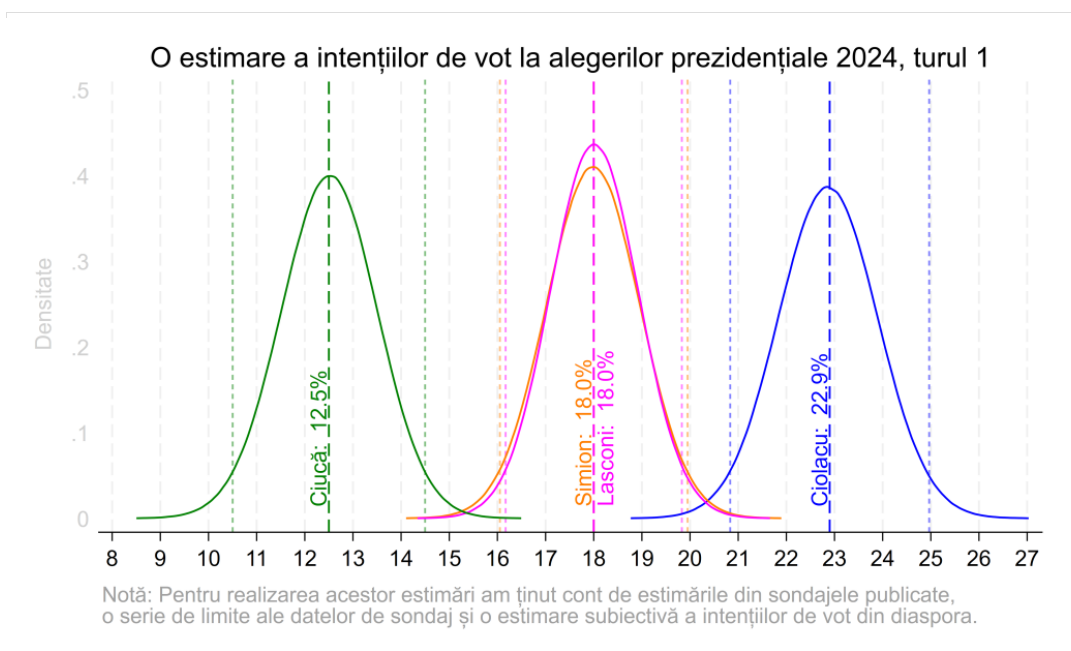
Să nu uităm de Verifield. Și acest institut a publicat ieri un sondaj (l-am inclus în estimările de ieri). În cazul lor nu prea sunt informații publice. Am găsit totuși două sondaje recente pentru

care știm rezultatul final, ambele realizate cu ocazia alegerilor pentru funcția de primar general din acest an. În primul sondaj, realizat cu trei luni înainte de alegeri, au estimat total greșit câștigătorul (vezi [aici](#)). În celălalt, realizat cu aproape două luni înainte, l-au estimat foarte bine (vezi [aici](#)). Prin urmare, nici în cazul lor nu știm dacă estimările sunt de încredere, respectiv cât de bine le va funcționa metodologia în cazul alegerilor naționale.

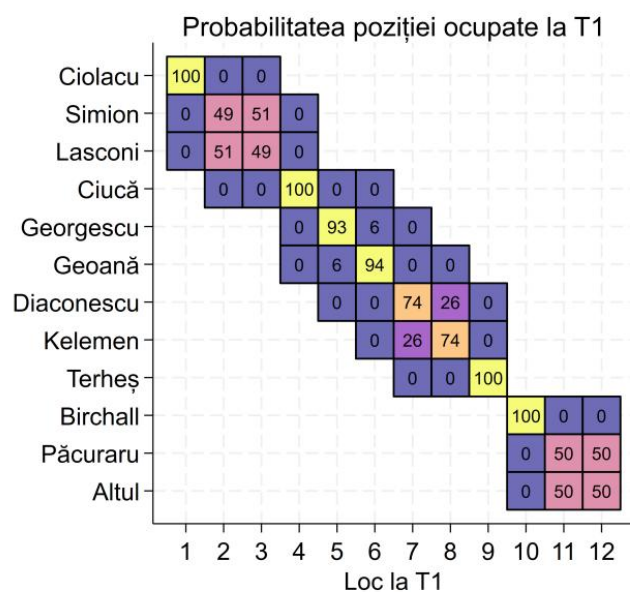
După aflarea rezultatelor de la aceste alegeri vom avea mai multe informații, prin urmare vom putea evalua performanța noilor metodologii (colectarea datelor, eșantionare, corecții). Până atunci, DUBITO!

Revenind la alegerile prezidențiale, dincolo de niște valori pur simbolice, nu s-a schimbat nimic semnificativ față de ieri: Simion și Lasconi sunt tot la egalitate, deși egalitatea de azi „e ceva mai egală.” Prin urmare, postez doar graficele cu estimările rezultate în urma modelării. Reamintesc faptul că **estimările de ieri și de azi se referă la rezultatul final**, deci țin cont și de voturile din diaspora.





Cu ocazia alegerilor prezidențiale din Moldova am realizat o serie de stimări ale intențiilor de vot, calculând și probabilitatea ca un candidat să câștige alegerile (vezi [aici](#), [aici](#) și [aici](#)). Fac ceva similar și pentru prezidențialele din România. Pentru că, în acest caz, sunt mai multe valori de afișat, prezint datele într-o formă grafică diferită. Valorile din grafic indică, pentru fiecare candidat, probabilitatea de a ocupa un anumit loc la finalul competiției. Dacă estimările din sondaje și modelarea sunt corecte, Ciolacu se va situa pe primul loc, Lasconi și Simion vor ocupa locurile doi sau trei, iar Ciucă va fi pe locul patru.



Prezidențiale 2024, T1: dinamica ratei de prezență

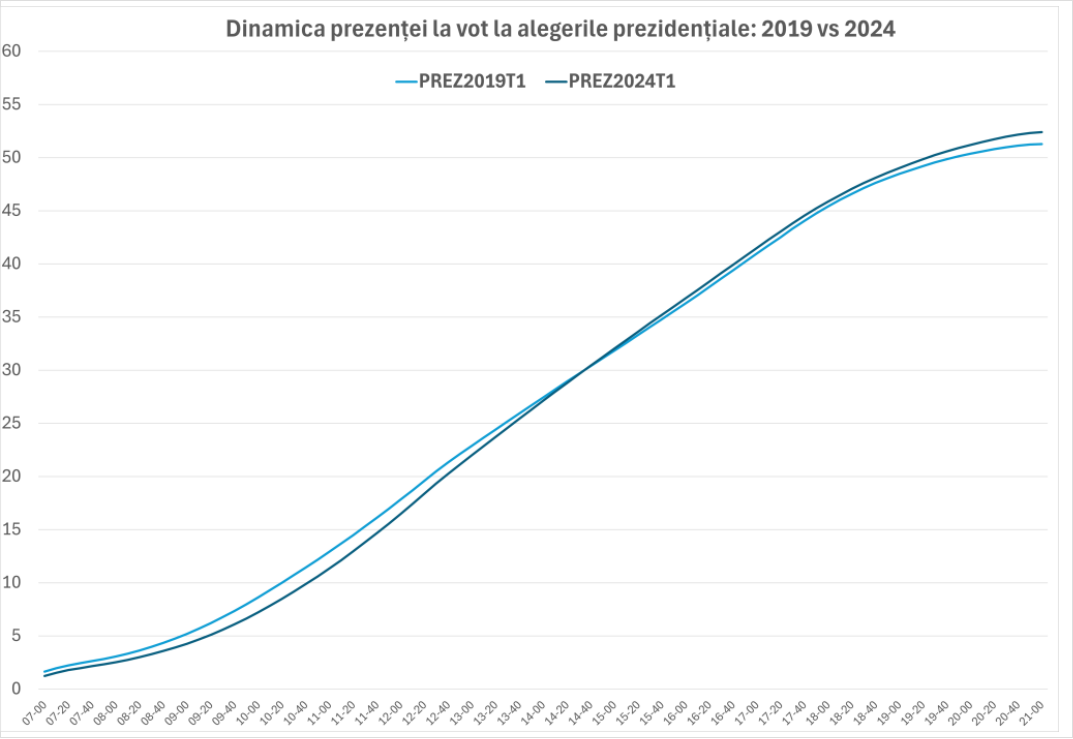
Data: 2024-11-24

Cuvinte cheie: alegerile-prezidentiale-2024, rata-de-prezenta-la-vot

Rata de prezență la interval de 10 minute, orele 7-17. Voi schimba graficul aici, din când în când.

Inițial, o rată de prezență mai redusă azi comparativ cu 2019. Diferența tinde să se micșoreze ușor. Da, s-au întâlnit pe la 15! La 17 prezența e deja puțin mai mare decât în 2019. La 19 nimic nou. Nimic nou în continuare.

Cel mai probabil, rata de prezență finală va fi în jur de 53%. Valorile din grafic sunt mai mici comparativ cu valorile oficiale cu aproximativ un punct procentual (unele date intră mai târziu în sistemul STS).



Prezidențiale 2024, T1: rata de participare în funcție de vârstă

Data: 2024-11-25

Cuvinte cheie: alegerile-prezidentiale-2024, rata-de-prezentat-la-vot

În această postare e vorba de „adevărata” **rată de participare**, nu de **structura votanților în funcție de vârstă**. Uneori, aceasta din urmă este greșit prezentată ca rată de prezență.

Datele de pe site-ul roaep prezintă numărul votanților de diferite vârste. Pornind de la valorile absolute, în unele cazuri au fost publicate grafice care indică ponderea diferitelor categorii de vârstă din total votanți. De exemplu, putem afla că au votat 10 milioane de persoane din care 1,3 milioane au fost tineri. Folosind aceste valori, putem calcula care este ponderea votanților tineri în totalul votanților (13% pentru acest exemplu). Strict în baza acestei informații, nu este corect să afirmăm că tinerii (nu) votează sau că votează într-o măsură mai mare/mică comparativ cu alte categorii de vârstă.

Cu siguranță, informația cu privire la numărul votanților de diferite vârste este foarte importantă. Însă, nu este singura importantă sau neapărat cea mai importantă. Depinde de ce ne interesează să aflăm. Pe baza datelor din Registrul electoral, sistemul STS ar fi putut să ofere aceleași informații și cu privire la alegători, nu doar la votanți. Sau, măcar să fi calculat și afișat **ratele de participare în interiorul fiecărei categorii de vârstă (gen)**. Am fi putut astfel compara gradul de participare la vot (rata

de prezență) în funcție de vârstă (gen). Astfel, am putea răspunde la întrebări de tipul:

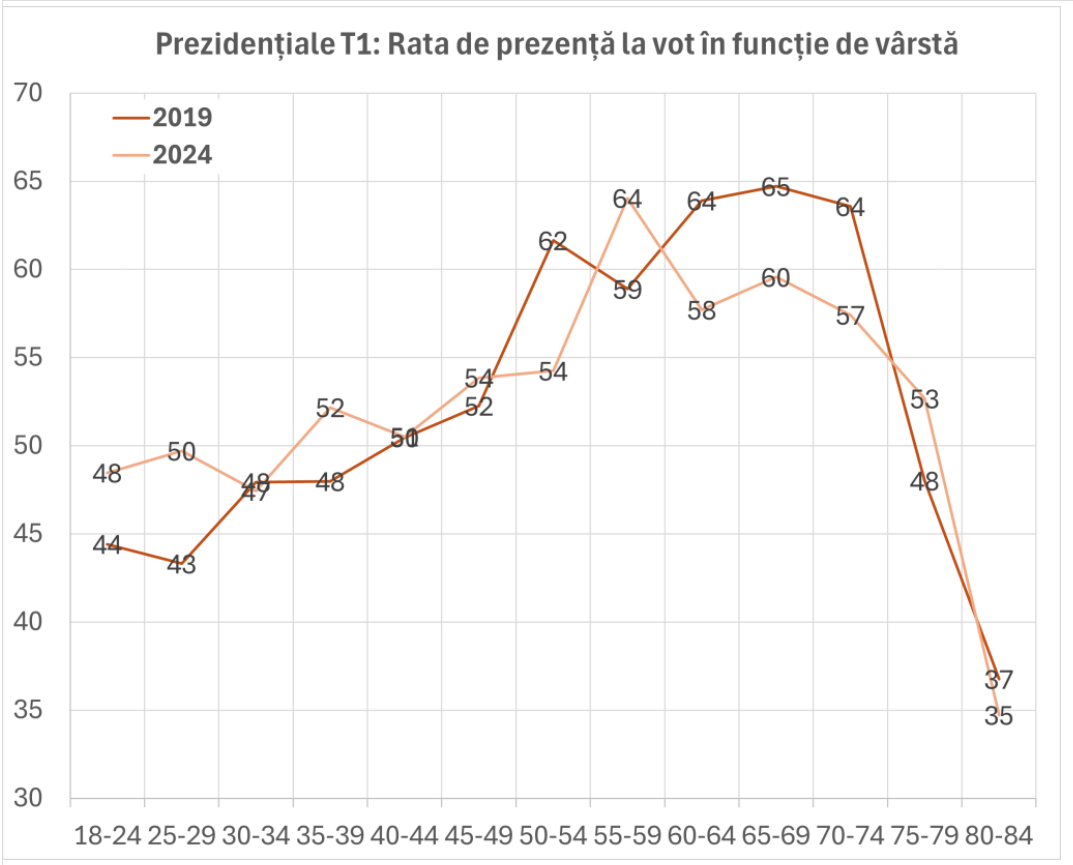
- Ce categorie de vârstă a avut cea mai mare rată de prezență / participare?
- Cum a variat rata de prezență a tinerilor de la un ciclu electoral la altul?

Cu puțin efort, am calculat ratele de prezență în funcție de vârstă folosind date de pe Tempo-Online (populația rezidentă pe categorii de vârstă), date Eurostat (emigranții pe categorii de vârstă) și, evident, numărul votanților conform roaep. Prin însumarea primelor două valori am ajuns la o valoare destul de apropiată de numărul alegătorilor de pe listele permanente așa cum apare pe roaep (18 milioane). Datele prezentate sunt calculate ținând cont de alegătorii, respectiv votanții din țară și diaspora (în cazul diasporei datele despre emigranți acoperă doar țările din Europa, deci reprezintă o aproximare în jos).

În cazul primului tur al alegerilor prezidențiale din 2019 și 2024, ratele de prezență în funcție de vârstă, urmează o curbă în formă de clopot (ceea ce nu e deloc o surpriză). Diferențele dintre cele două competiții sunt în general mici (chiar dacă e posibil să fie mici probleme cu datele folosite).

Pentru a ușura lectura, prezint două exemple de citire:

- 48% dintre tineri (18-24 ani) au votat la aceste alegeri, respectiv 44% dintre tineri (18-24 ani) au votat în 2019. Dacă îi excludem pe cei de peste 80 de ani, tinerii (18-24) au cea mai scăzută rată de prezență la vot (alături de cei de 25-29 și 30-34 ani).
- Comparativ cu turul 1 al alegerilor prezidențiale din 2019, în 2024 tinerii (18-29) s-au prezentat la vot într-o pondere mai mare. În cazul seniorilor (60-74), situația stă exact invers, rata de participare la vot fiind mai mare în 2019 comparativ cu 2024.



Un altfel de grafic

Data: 2024-11-25

Cuvinte cheie: alegerile-prezidentiale-2024, dinamica-intentiilor-de-vot

Pun aici o imagine de cu totul altă factură decât cele postate anterior. Alături de schimbarea de dinamica de ultim moment pe care am semnalat-o sâmbătă în postarea cu titlul [Călin file de poveste](#), aceasta imagine ne-ar putea ajuta să înțelegem cumva din interior surpriza de la turul 1.

Nimic din ce apare în imagine nu are legătură cu ceea ce cred sau nu cred. Tot ceea ce am făcut a fost să pun alături câteva frânturi de povești din mediul online. Mă întreb dacă nu am uitat o dimensiune importantă? Mai mult ca sigur că da. Poate data viitoare voi face mai „serios” o astfel de analiză, cu date cantitative și eșantioane reprezentative ale conținuturilor, nu cu impresii marginale, produse rapid. Oare acum mai suntem la fel de surprinși de rezultatul de ieri?!



P.S. Azi mi-am făcut cont pe TikTok. Am vrut să fac asta mai demult, pentru un cu totul alt motiv (cine are copii, mă înțelege). Sper să-l folosesc cu moderație și strict în interes științific.

Parlamentare 2024: o estimare a intențiilor de vot

Data: 2024-11-29

Cuvinte cheie: alegerile-parlamentare-2024, modelarea-intențiilor-de-vot

Reiau exercitiul de predicție a rezultatelor alegerilor pe baza combinării estimărilor produse de sondaje (pentru prezidențiale, vezi [aici](#) și [aici](#)). De această dată e rândul alegerilor parlamentare. Nici acum nu am avut acces la micro-date (nu că nu aș fi vrut și/sau nu ar fi fost mai bine). Și în acest caz, **predicția se referă la rezultatul final**, cel care include voturile din țară plus diaspora. Informațiile utilizate pentru producerea estimărilor intențiilor de vot la parlamentare au fost următoarele:

- **estimări ale intențiilor de vot conform sondajelor** disponibile în spațiul public; aceasta a fost principala sursă de date brute; din păcate, după turul 1, au fost publicate doar două sondaje; în cazul acestor date am aplicat mici corecții; modificările au fost în medie de un punct procentual, în sus sau în jos, funcție de direcția în care am apreciat că vor acționa diferiți factori (rată răspuns, dezirabilitatea socială);
- **modelări ale datelor de sondaj**; în principal estimarea tendințelor generale și, mai ales, a tendințelor din ultimele două săptămâni;
- **numărul voturilor primite de candidați la turul 1 al alegerilor prezidențiale**;

- **matricea de transfer a voturilor de la prezidențiale (turul 1) la parlamentare;** am plecat de la numărul voturilor conform datelor oficiale și le-am ponderat conform acestei matrici; pentru estimarea valorilor de input am luat în calcul afinitățile ideologice dintre votanții candidaților și partide, coroborate cu informația din matricile calculate pe baza datelor din trei sondaje independente; mulțumesc [IRES](#), [Agenția de Rating Politic](#) și [AtlasIntel](#) pentru furnizarea acestor date; desigur, am ținut cont și de faptul că, foarte probabil, nu toți alegătorii care au votat la turul 1 vor vota și la alegerile parlamentare.

Principalele „puncte slabe” ale metodei: calitatea datelor de sondaj, numărul foarte mic al sondajelor recente (doar două), numărul mic al respondenților, în special în cazul partidelor ceva mai mici, estimarea matricii de transferuri, mai ales în ceea ce privește (1) transferul voturilor candidatului Călin Georgescu spre partide și absenteiști și (2) numărul total al votanților la alegerile parlamentare.

Suplimentar, în cazul partidelor mici, (1) **distribuția socio-spațială a votanților** este foarte eterogenă, iar (2) **votul tactic** va fi prezent într-o măsură ceva mai mare de această dată, mai ales dacă aceste partide se adresează unui electorat mai educat și mai rațional. Un astfel de electorat va cumpăni atent și va acorda votul nu doar în funcție de orientarea ideologică a unui partid ci și în funcție de șansele ca acesta să treacă sau nu pragul electoral. De altfel, votul tactic a funcționat și în cazul primului tur de la prezidențiale, avantajând-o pe Elena Lasconi.

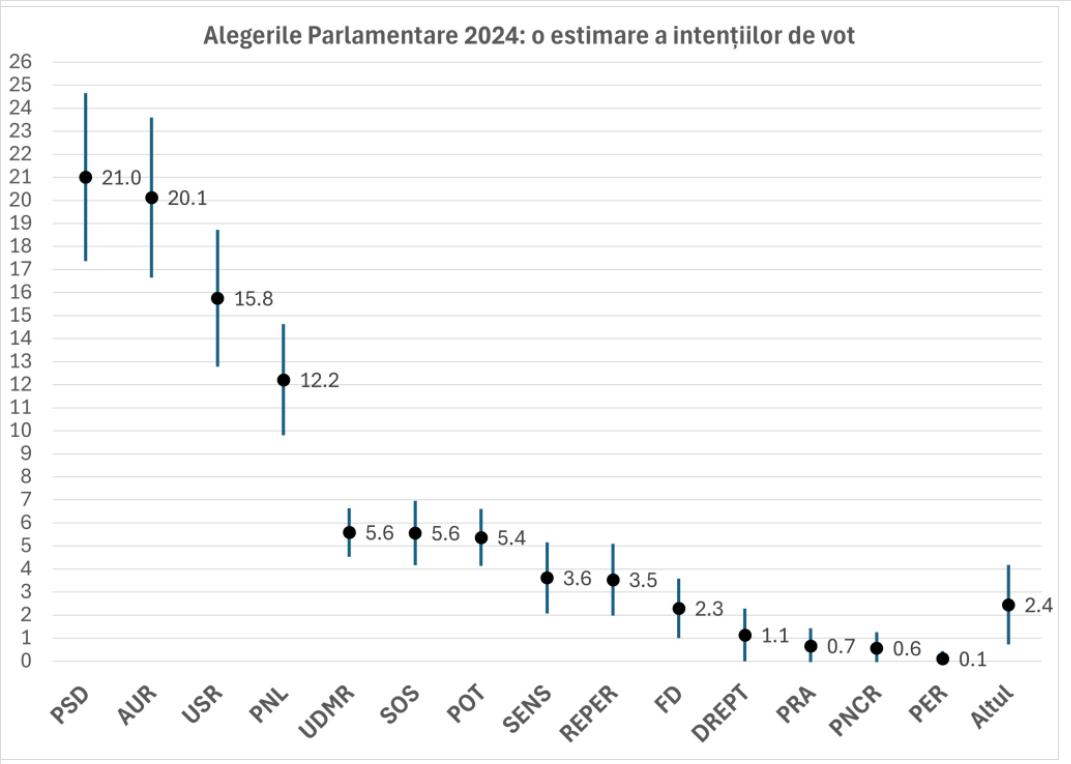
Toți acești factori contribuie la **creșterea incertitudinii estimărilor**. Dacă, în cazul partidelor mai mari, o diferență de unul-două procente nu contează prea mult, în cazul partidelor mici, aceeași diferență poate avea **consecințe majore**: partidul respectiv poate trece sau nu pragul electoral (5%), deci poate intra sau nu în Parlament.

Similar cu estimările făcute în cazul alegerilor prezidențiale, și în acest caz **am lucrat cu distribuții probabile ale estimărilor**, nu cu valori fixe. Pentru alte informații cu privire la metoda folosită se

pot citi postările de [aici](#) și [aici](#). Prezint rezultatul modelării într-o formă grafică simplă: cercurile pline indică estimările intențiilor de vot, iar liniile indică incertitudinea estimărilor, adică intervalul în care ne așteptăm, cu o probabilitate de 95%, să se situeze rezultatul final.

Pe scurt, mesajul acestui grafic e următorul:

- **primul loc este disputat de PSD și AUR;** în cazul acestor partide, distribuțiile estimărilor se suprapun în mare măsură, cu un ușor plus pentru PSD (21% vs. 20%); procentele finale depind de gradul de mobilizare a activului PSD, respectiv de numărul votanților lui Georgescu care se vor prezenta la urne și vor vota cu AUR.
- **pe locul trei se situează USR** cu aproximativ 16%; scorul final depinde de numărul voturilor tactice (dinspre și/sau înspre REPER și SENS); funcție de cum se vor muta aceste voturi, polul progresist poate obține un scor final foarte diferit și, în consecință, un număr mai mic sau mai mare de parlamentari;
- **pe locul patru se situează PNL** cu aproximativ 12%; scorul final depinde de impactul venirii la conducerea PNL a lui Bolojan, respectiv de acordul alegătorilor cu reformele propuse de acesta;
- în continuare, în zona imediat superioară pragului electoral (5%), se situează partidele **SOS, UDMR și POT**; deși intervalul de variație al estimărilor acoperă parțial și zona de sub 5%, destul de probabil aceste partide vor trece pragul electoral;
- **REPER și SENS se situează imediat sub pragul electoral;** funcție de cât de mare va fi mobilizarea și de cum va arăta distribuția voturilor în interiorul polului progresist, aceste partide vor trece sau nu pragul;
- restul partidelor sunt destul de mult sub prag; dintre acestea, FD are totuși ceva șanse de a intra în Parlament, dar e nevoie de o mobilizare excepțională.

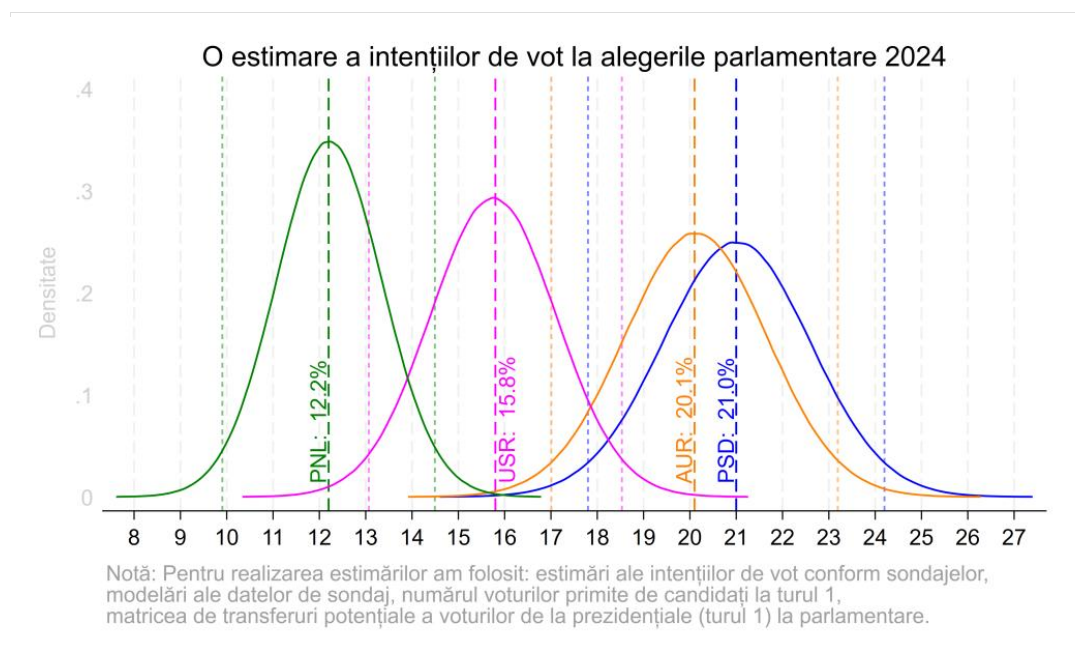


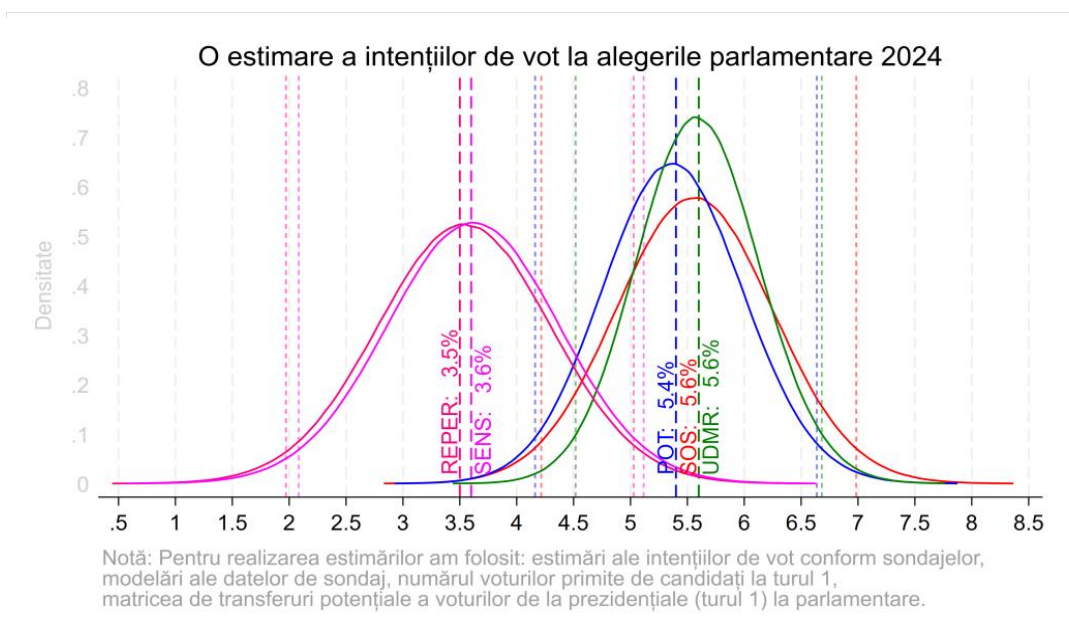
... tot Parlamentare, ceva mai fancy

Data: 2024-11-30

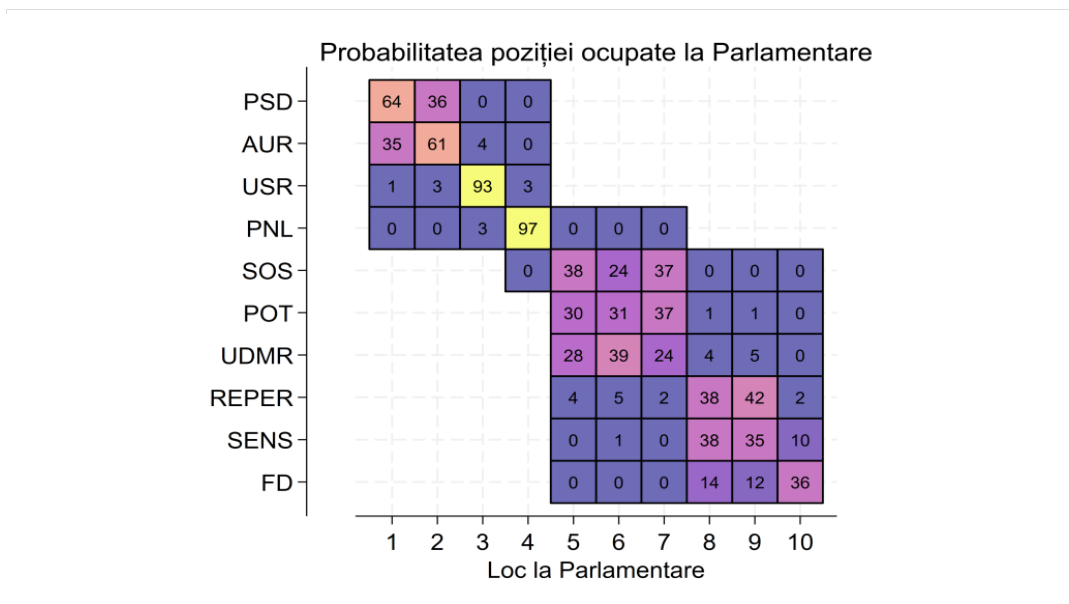
Cuvinte cheie: alegerile-parlamentare-2024, modelarea-intentiilor-de-vot

Estimarea din [postarea anterioară](#), în haine noi, de duminică! Pentru un exemplu similar, cu interpretare, vezi [aici](#). Merită să ne reamintim estimarea de la turul 1, nu? Cine uită, riscă să ...! Atunci, în ultima săptămână, au fost publicate două sondaje. Trendul a devenit vizibil târziu și a fost exponențial, după cum am sugerat, pe baza unor date limitate, [aici](#). Acum ne-a fost „servit” doar un sondaj (realizat online de AtlasIntel), hai două (tot online, dar dispărut între timp). Prin urmare, „Om trăi și om vedea!”





Adaug un grafic cu probabilitățile asociate estimărilor, dar doar pentru primele 10 locuri în ordinea descrescătoare a procentelor probabile. PSD are cele mai mari șanse să ocupe primul loc (64%), AUR locul secund (61%) (și 35% primul loc), USR locul trei (93%) ș.a.m.d. Desigur, șansele ca realitatea să arate diferit nu pot fi ignorate: AUR pe primul loc (35%) și PSD pe doi (36%).

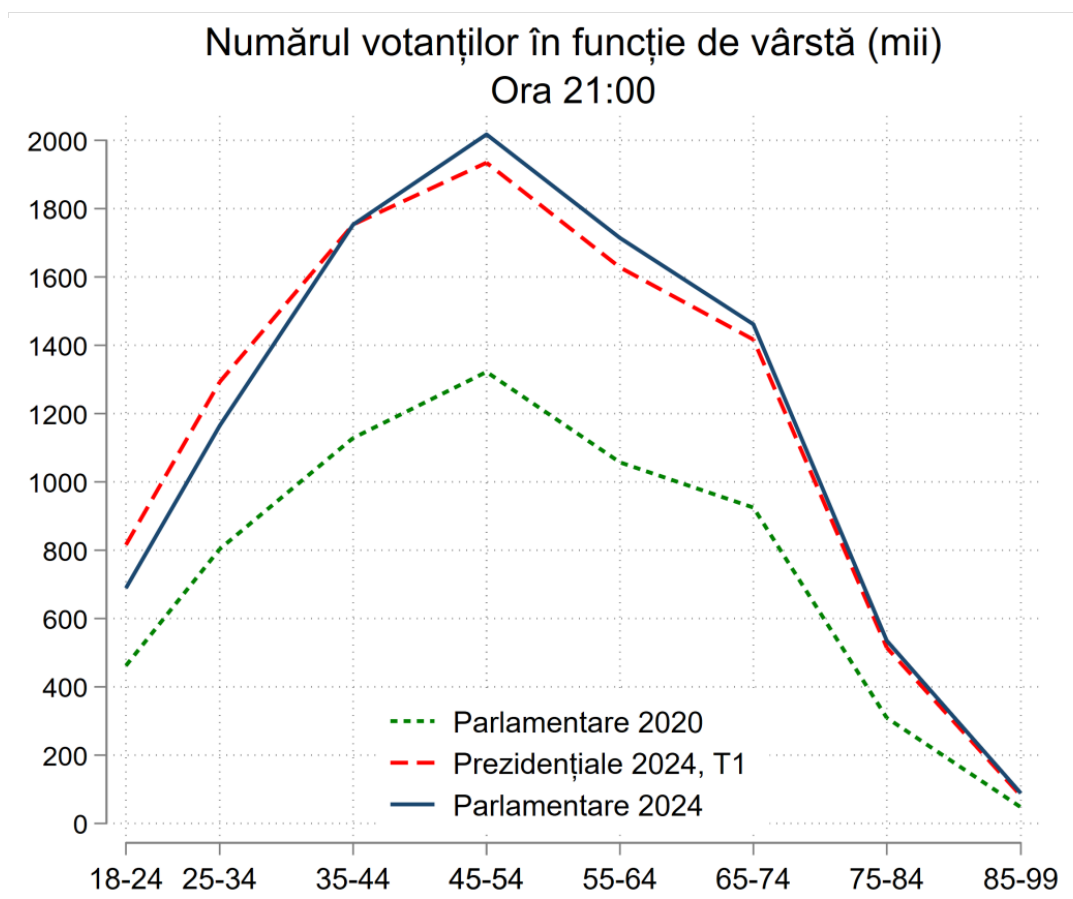


Numărul votanților în funcție de vârstă / anul nașterii

Data: 2024-12-01

Cuvinte cheie: alegerile-parlamentare-2024, numarul-votantilor-in-functie-de-varsta

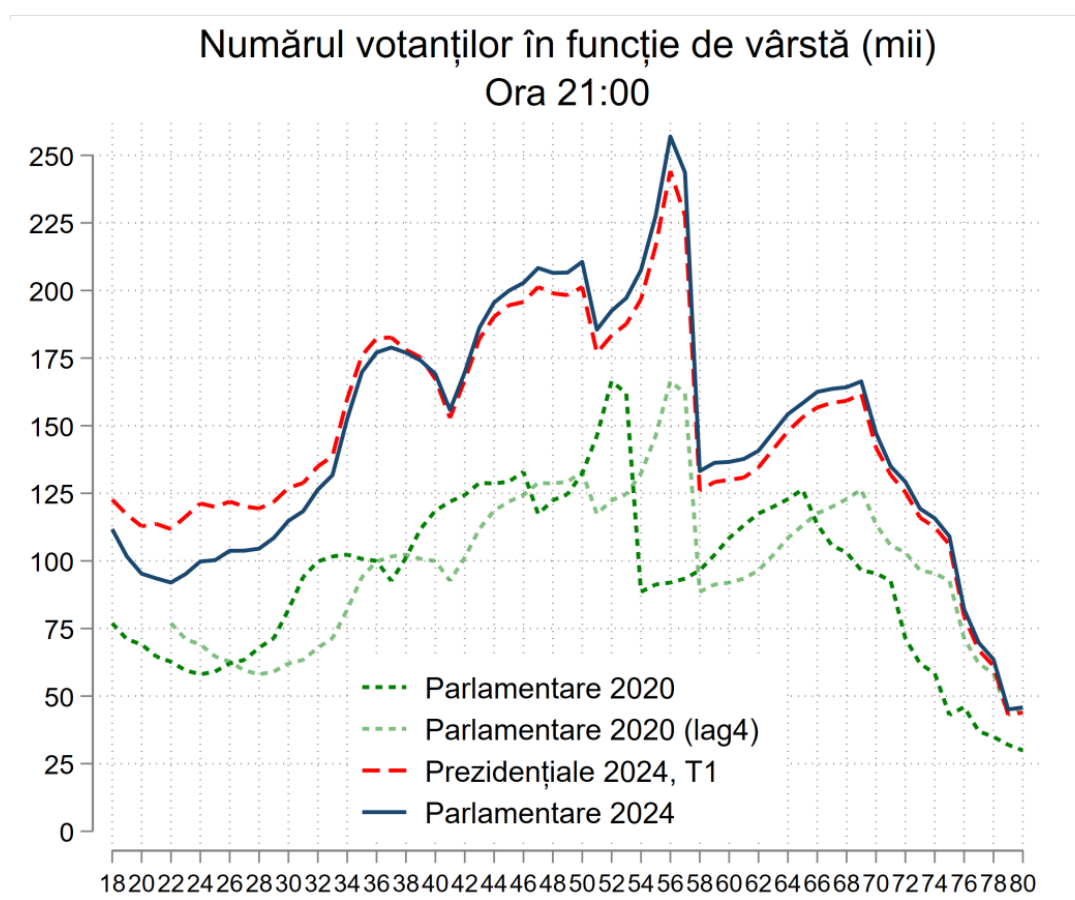
Pun aici, pe parcursul zilei, un grafic cu numărul votanților pe categorii de vârstă la alegerile Parlamentare 2024, comparativ cu Parlamentare 2020 și Prezidențiale 2024, T1.



Pentru cei interesați de detalii, pun și graficul pe vârste. Pentru că alegerile parlamentare 2020 au avut loc în urmă cu patru ani, am adăugat o linie frântă care descrie un scenariu ipotetic: care ar fi fost în 2024 numărul votanților de la Parlamentare 2020 într-un context cu COVID. Câteva concluzii:

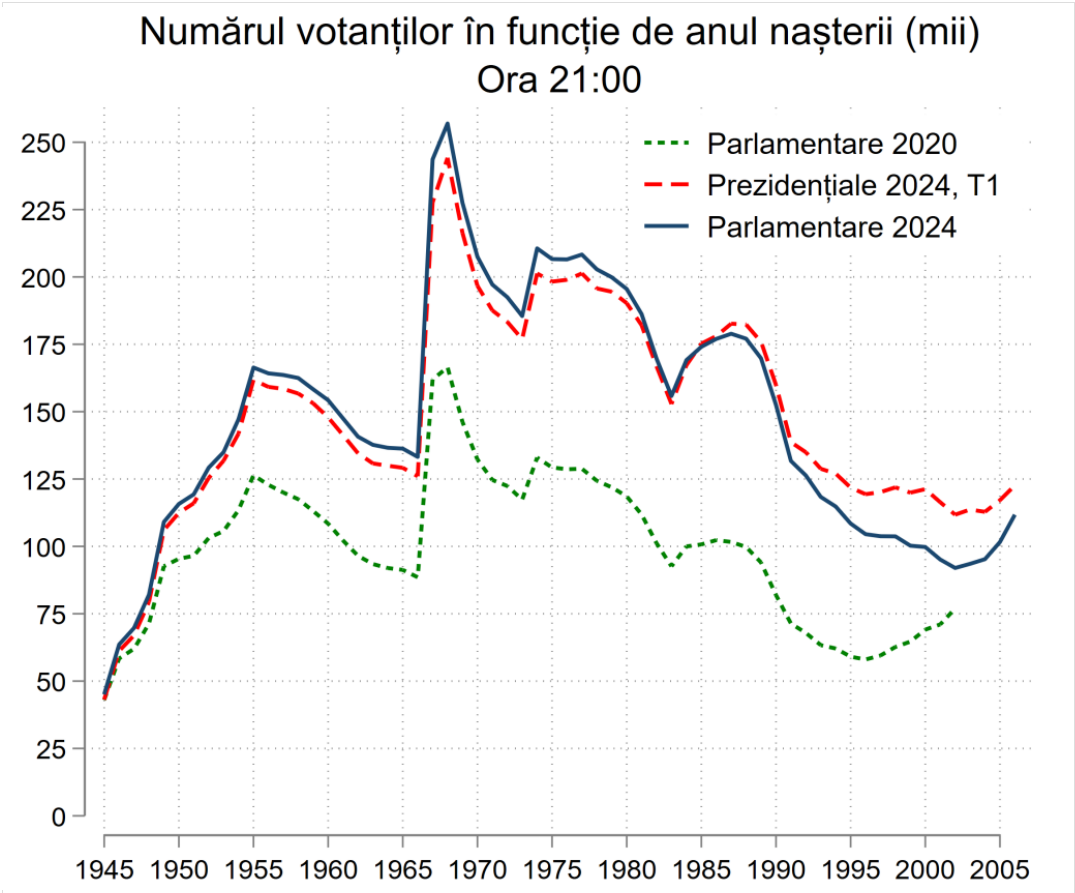
- Linia frântă de la Parlamentare 2024 este similară cu cea de la Prezidențiale 2024 (T1), deci numărul votanților urmează același patern.
- Comparativ cu Prezidențiale 2024 T1, numărul celor care votează la Parlamentare 2024 este mai mare, cu excepția tinerilor (deocamdată).

- În cazul alegerilor parlamentare din 2020 se observă impactul COVID-ului asupra prezenței la vot, indiferent de vârstă (cu unele variații).
- Se observă un decalaj de aproximativ patru ani între curba de la Parlamentare 2020 și cea de la Parlamentare 2024. Pentru o comparație corectă, curba de la Parlamentare 2020 trebuie translatată la dreapta cu patru ani.
- Diferențele se văd mai bine atunci când ne raportăm la anul nașterii (ultimul grafic).



În toate aceste grafice este vorba despre numărul votanților. Pentru că numărul de persoane de o anumită vârstă, respectiv dintr-o anumită generație diferă, nu putem spune că o vârstă /

generație votează într-o măsură mai mare sau mai mică. Putem face astfel de comparații doar dacă știm rata de prezență la vot. Va urma!

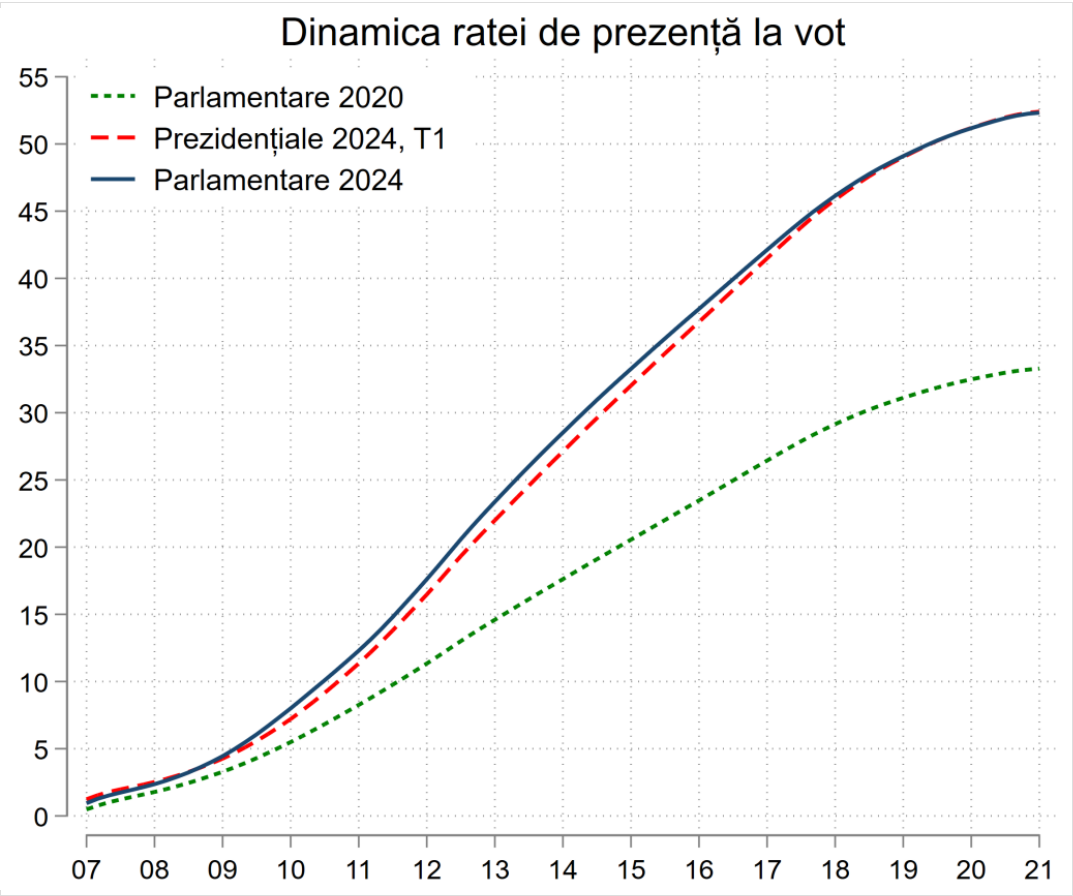


Parlamentare 2024: dinamica ratei de prezență

Data: 2024-12-01

Cuvinte cheie: alegerile-parlamentare-2024, rata-de-prezenta-la-vot

Pun aici, pe parcursul zilei, rata de prezență la interval de 10 minute. Cum lucrurile capătă mai mult sens atunci când sunt puse în context, am inclus și dinamica prezenței de la Parlamentare 2020, respectiv Prezidențiale 2024, T1.

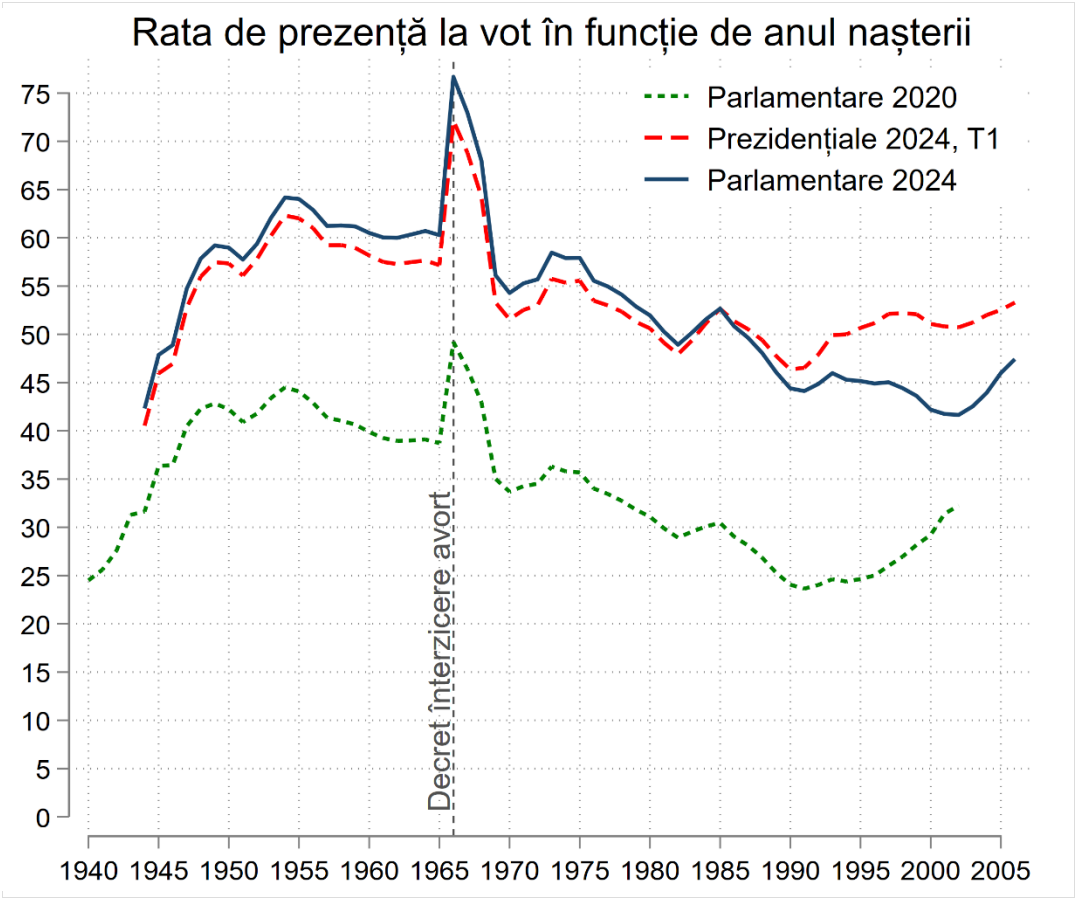


Rata de prezență la vot în funcție de anul nașterii

Data: 2024-12-03

Cuvinte cheie: alegerile-parlamentare-2024, alegerile-prezidentiale-2024, rata-de-prezenta-la-vot

Graficul promis, cu câteva precizări. Evident, ar fi fost mai simplu ca STS să ofere și numărul persoanelor din listele electorale pentru fiecare combinație de vârstă și sex. Cum STS nu a făcut asta, am folosit indicatorul „Populația după domiciliu la data de 1 ianuarie” (POP107A, Tempo-Online). Alternativ, puteam să folosesc populația rezidentă, dar, pentru că am inclus și votanții din diaspora, populația cu domiciliul e oarecum mai potrivită în acest caz (Eurostat nu are date despre emigranți pe intervale de vârstă de un an). Pentru că datele din cele două surse (STS și Tempo-Online) au o dată de referință diferită, valorile din grafic sunt medii mobile (valoarea din anul anterior, anul de referință și anul următor).



Dacă „sondaje nu e, nimic nu e!” Sau ...

Data: 2024-12-05

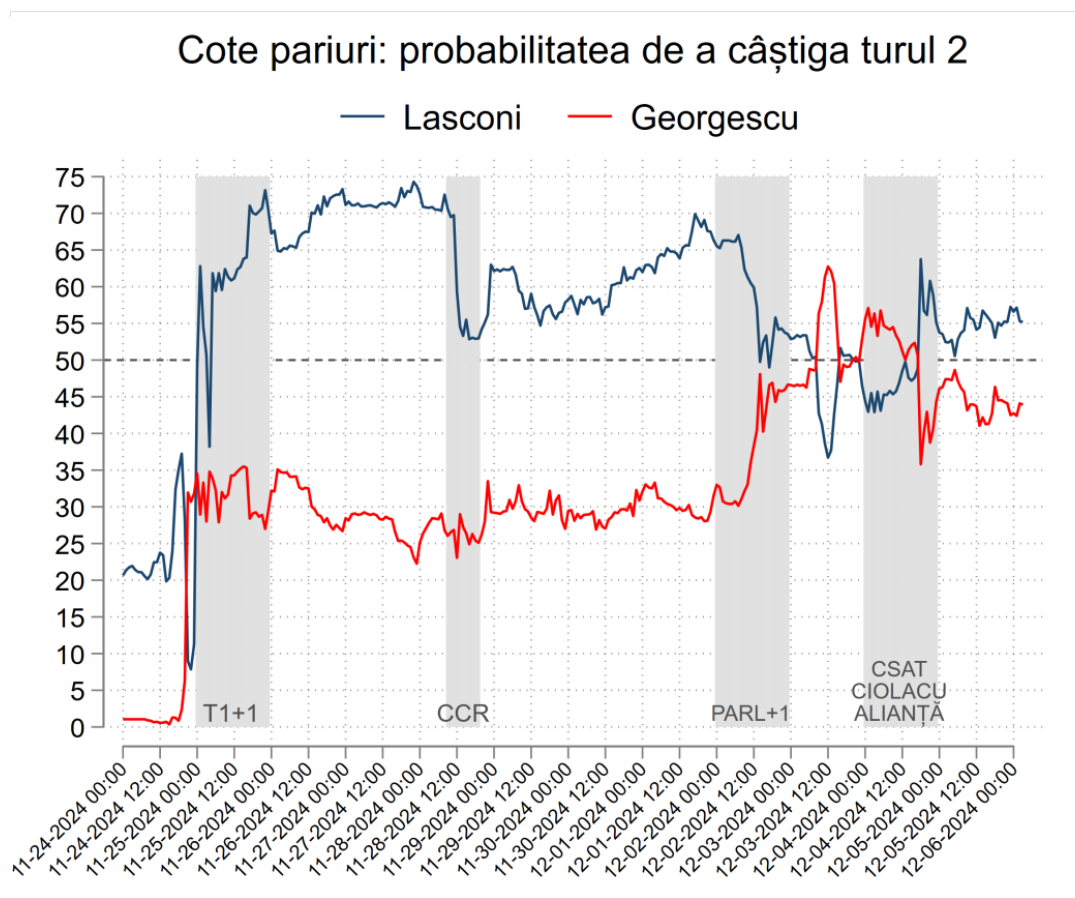
Cuvinte cheie: alegerile-prezidentiale-2024

Săptămâna aceasta au apărut foarte puține date de sondaj: exit-poll-ul CURS de duminică și sondajul AtlasIntel de ieri (cu date colectate în perioada 2-4 decembrie). Niciunul dintre aceste sondaje nu a acoperit și seara zilei de 4 decembrie, seară marcată de trei știri importante: (1) desecretizarea documentelor CSAT privind alegerile, (2) declarația premierului Ciolacu cu privire la susținerea candidatei Elena Lasconi de către PSD și (3) demararea procesului de constituire a unei majorități parlamentare și guvernamentale pro-europene formată din PSD, PNL, USR și UDMR.

În lipsa unor date de sondaj e dificil de estimat efectul acestor decizii asupra intențiilor de vot de la turul 2. Desigur, e mai probabil ca efectul să fie unul pozitiv pentru Elena Lasconi. Cotele de la casele de pariuri ne pot oferi o estimare aproximativă (totuși, probabil distorsionată). E de preferat ca volumul pariurilor să fie cât mai mare, iar datele să fie publice. Am găsit o sursă potrivită, am descărcat datele, le-am aranjat nițel și *voilà* minunăție!

Pentru a ghida interpretarea, am marcat pe grafic patru evenimente importante: prima zi de după turul 1 (rezultatul oficial al alegerilor), decizia CCR de renumărare a voturilor de la turul 1, prima zi de după alegerile parlamentare (rezultatul oficial al alegerilor) și ziua de 5 decembrie. Oare cum se poate explica scăderea bruscă a cotei Elenei Lasconi din data de 3 decembrie, înainte de amiază? Foarte probabil e efectul declarațiilor liderilor

PSD, în principal ale premierului Marcel Ciolacu care a afirmat că va lăsa alegătorii PSD să decidă singuri cum vor vota la turul 2.



Atenție, cotele caselor de pariuri nu reprezintă o predicție a rezultatului de la turul 2. Cei care pariază nu constituie un eșantion reprezentativ de votanți, iar modul în care sunt stabilite cotele este influențat și de alți factori decât preferințele exprimate de cei care pariază. Cu toate acestea, unele modele de predicție iau în considerare și cotele caselor de pariuri.

Prezidențiale 2024, T2: o estimare a intențiilor de vot

Data: 2024-12-06

Cuvinte cheie: alegerile-prezidentiale-2024, modelarea-intențiilor-de-vot

Deși nu mai e de actualitate, continui exercițiul de predicție a rezultatelor alegerilor pe baza combinării estimărilor produse de sondaje (pentru alegerile prezidențiale T1, vezi [aici](#) și [aici](#)). Nici acum nu am avut acces la micro-date (nu că nu aș fi vrut și/sau nu ar fi fost utile). Și în acest caz, **predicția se referă la rezultatul final, cel care include voturile din țară plus diaspora**. Informațiile utilizate pentru producerea estimărilor intențiilor de vot la turul 2 al alegerilor prezidențiale au fost următoarele:

- **estimări ale intențiilor de vot conform sondajelor;** aceasta a fost principala sursă de date brute; am folosit estimările din sondajele publicate realizate după turul 1 al alegerilor prezidențiale; am identificat opt sondaje, din care patru au fost realizate între turul 1 și parlamentare, două în ziua alegerilor parlamentare și două după alegerile parlamentare (ambele realizate de AtlasIntel); dintre ultimele două sondaje, doar unul a prins perioada de după 4 decembrie; dacă vă întrebați de ce data de 4 decembrie e importantă, citiți vă rog postarea de [aici](#); de această dată, **consensul cu privire la estimări a fost și mai redus ca de obicei**; desigur, sondajele au fost realizate în zile diferite, au folosit tehnici diferite, deci era de așteptat să observăm o variabilitate mare a estimărilor; totuși, o diferență de 17

puncte procentuale între estimările produse de două sondaje realizate în ziua alegerilor parlamentare e mult prea mult; informațiile cu privire la modul în care au fost produse estimările din sondaje lipsesc cu desăvârșire (deocamdată nu sunt publicate nici măcar rapoartele obișnuite, unele de altfel minimale); **am asumat că e nevoie de mici corecții ale acestor estimări**; modificările au fost de cel mult un punct procentual, în sus sau în jos, în funcție de direcția în care am apreciat că vor acționa diferiți factori (rată răspuns, dezirabilitatea socială, tehnica de anchetă);

- **modelarea datelor de sondaj**; în acest caz am încercat să identific tendințele, mai ales cele asociate datei de 4 decembrie; din păcate datele publice care să includă și ziua de 5 decembrie sunt minime și provin de la un singur institut (AtlasIntel) (pot cel mult visa la date colectate azi); de asemenea, ar fi fost util să includ estimările produse de un sondaj recent realizat printr-o tehnică de anchetă diferită de cea folosită de AtlasIntel;
- **numărul voturilor primite de candidați la turul 1 al alegerilor prezidențiale, respectiv numărul voturilor primite de partide la alegerile parlamentare**;
- **matricea de transfer a voturilor de la turul 1 la turul 2, respectiv de la parlamentare la turul 2**; aici am avut două surse de date (AtlasIntel și CURS; mulțumiri ambelor institute); în ambele cazuri, matricile sunt anterioare zilei de 4 decembrie (1 decembrie și 28 noiembrie), deci nu prind impactul evenimentelor recente; am plecat de la numărul voturilor conform datelor oficiale și le-am ponderat conform acestor matrici;
- am încercat să țin cont și de faptul că, foarte probabil, vom asista la **o modificare a prezenței la vot**: (1) vor fi alegători care au votat la turul 1 și/sau parlamentare și care nu vor vota la turul 2 și (2) **vor fi alegători care anterior nu au votat, dar vor vota la turul 2**; numărul votanților noi va fi mult mai mare; cât de mare, e dificil

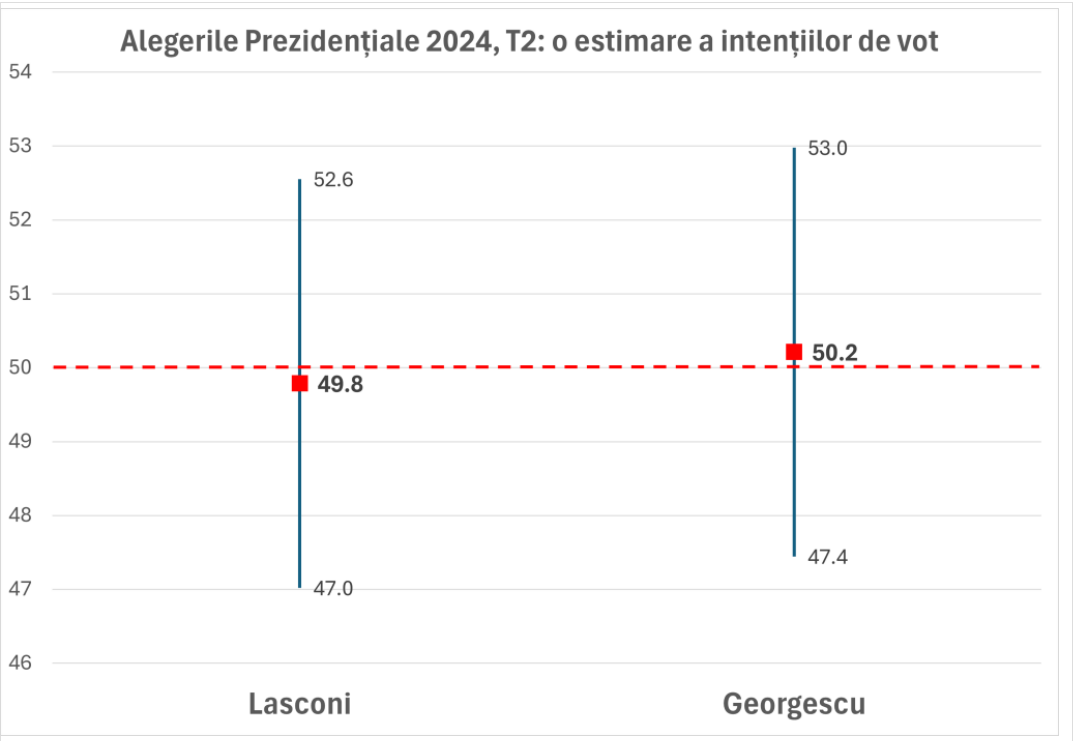
de spus; dacă ar fi să ghicesc, aş zice 1-2 milioane, deci vor fi aproximativ 10,5-11 milioane de votanţi; la fel de important e şi **cum vor vota votanţii noi** / care va fi distribuţia voturilor / cât de mare va fi diferenţa dintre numărul de voturi noi primite de fiecare finalist; **calitatea modelării depinde în principal de răspunsul la întrebarea: "Câţi votanţi noi va aduce la urne fiecare dintre cei doi candidaţi?"**;

- o ultimă componentă este **probabilitatea de câştiga alegerile conform caselor de pariuri (cotele)**; am scris despre această sursă de date [ieri](#), aşa că nu mai reiau aici.

Principalele „puncte slabe” ale metodei: calitatea datelor de sondaj, existenţa unui singur sondaj recent, posibile erori asociate matricilor de transferuri, şi, în principal, **numărul şi distribuţia voturilor noi**, respectiv **variabilitatea mare a estimărilor produse de sondaje**. Toţi aceşti factori contribuie la **creşterea incertitudinii estimărilor**.

Am lucrat cu distribuţii probabile ale estimărilor, nu cu valori fixe. Pentru alte informaţii cu privire la metoda folosită se pot citi postările de [aici](#), [aici](#) şi [aici](#).

Prezint rezultatul modelării într-o formă grafică simplă: pătratele pline de culoare roşie indică estimările intenţiilor de vot; liniile verticale indică incertitudinea asociată estimărilor, adică intervalul în care ne aşteptăm, cu o probabilitate de 95%, să se situeze rezultatul final, presupunând că asumpţiile care au stat la baza fiecărei componente a modelării sunt corecte (sau măcar nu se îndepărtează prea mult de la realitate). Desigur, **în prezent nu mai suntem în situaţia în care putem valida acest demers de estimare**.



Opinie: De ce candidați independenți? De ce doi?

Data: 2024-12-12

Cuvinte cheie: politica

Aș fi preferat să-mi fundamentez opiniile și pe analiza unor date din sondaje și din social media. Cum nu e cazul, rămâne să mă bazez pe ceea ce se numește „ghicire educată” însoțită, sper, de raționamente corecte.

De ce candidați independenți?

Imaginea partidelor și a clasei politice este foarte proastă. Cu excepția celor înregimentati politic și a celor care beneficiază de diferite avantaje, gradul de susținere al actualelor partide și reprezentanților acestora este foarte redus (sub 10%). Da, pe ici pe colo mai sunt figuri care nasc speranțe, dar imaginea generală e proastă.

Ce s-a stricat în 35 de ani nu poate fi reparat în câteva luni. Dezamăgirile repetate pot fi îndreptate doar prin confirmări pozitive repetate.

Afilierile de partid funcționează mult mai slab în cazul alegerilor prezidențiale, iar votul pro sau contra unei persoane este semnificativ mai puțin fundamentat pe criterii raționale.

De ce doi candidați?

Pentru a nu antagoniza și mai mult spațiul social și politic. Am în vedere atât populația, cât și clasa politică. Deja mergem de ceva vreme în această direcție, deja nivelul de coeziune socială este redus. Dacă vom continua, ca societate, să mergem în această direcție, nu ne așteaptă deloc lucruri bune. Istoria ne oferă destule exemple în acest sens. Un spațiu competitiv dominat de doi actori, deci două opțiuni, va produce întotdeauna mai mult antagonism comparativ cu un spațiu în care performează trei sau mai mulți actori.

Pentru ca alegătorii să beneficieze de o paletă ceva mai extinsă de opțiuni. Criteriile de selecție folosite de partide sunt cel mai adesea opace, au legătură mai degrabă cu loialitatea față de cei aflați în funcții de conducere și mult mai puțin cu profesionalismul și caracterul persoanelor. Adesea, alegătorii au fost nemulțumiți de propunerile partidelor pentru diferite funcții. De-a lungul anilor au văzut în jurul lor multiple situații în care persoane fără o calificare potrivită (sau fără nicio calificare), adesea de o moralitate îndoielnică, ajung să ocupe funcții importante, bine plătite și/sau care îi pun în poziția în care pot obține diferite avantaje, adesea ilegale. Ca să nu mai vorbim de situațiile în care s-au inventat instituții și posturi pentru a răsplăti clientela de partid. Nemulțumirea este cu atât mai mare în cazul alegerilor prezidențiale. În astfel de contexte, inadecvarea candidaților propuși de partide devine și mai evidentă. În cazul ultimelor alegeri prezidențiale lucrurile au stat chiar mai rău decât de obicei. O mare parte a populației nu s-a regăsit în niciunul dintre candidați. Datele din sondaje arată că o pondere mai mare decât de obicei a electoratului nu a avut o preferință (cel puțin nu una exprimată și/sau clar definită) cu câteva zile înainte de alegeri. Într-un context de nemulțumire generalizată și difuză, plus incertitudine, preferințele alegătorilor pot fi manipulate mult mai ușor, ceea ce s-a și întâmplat. O singură variantă, și aceasta aleasă de alții, în „laborator,” nu are cum să fie percepută ca o alegere reală. De prea multe ori în ultimii ani populația a fost pusă în fața unor situații în care a simțit că nu poate alege sau chiar nu a putut alege (chiar dacă pentru motive îndreptățite

uneori). Într-un astfel de context social, „soluția unică” ar putea să producă efecte contrare celor dorite.

Necesitatea unei poziționări strategice. Spațiul politic este tot mai puțin dominat de axa stânga-dreapta. Încet, încet, au ajuns să domine poziționările pe alte axe: suveranistă, anti-corupție, anti-sistem, anti-partide cartel. E nevoie de voci din exterior care să adreseze echilibrat și credibil aceste teme. Și, mai ales, să răspundă fricilor generate de acestea. E puțin probabil ca un singur contra-candidat să poată acoperi credibil toate aceste zone. E și mai puțin probabil ca un profil unic să corespundă preferințelor exprimate de cât mai multe dintre segmentele electorale majore. În plus, existența a două variante permite o poziționare (de preferat naturală, dar poate fi și una direcționată) care să acopere, cumulat, o zonă mai mare din preferințele alegătorilor. Într-un astfel de context, sunt șanse rezonabile ca noile opțiuni să domine spațiul dezbaterilor, să legitimizeze o altfel de imagine ideală de președinte, și, de ce nu, să ocupe primele două poziții la nivelul opțiunilor de vot. Ar fi prea frumos, nu? Riscant? Nu cred. Oricum, în cel mai rău caz, unul dintre cei doi candidați va ajunge în turul doi, deci nu se pierde nimic. Dacă există șanse ca doi candidați pro-europeni să ajungă în turul doi, iar șansele de a pierde ceva sunt minime, decizia rațională este evidentă, cred. Desigur, scenariul are sens doar dacă acești candidați se poziționează aproximativ la fel pe temele majore, respectiv nu se atacă (grav) unul pe altul în campanie. Dat fiind procesul prin care ar urma să fie aleși cei doi candidați și profilul așteptat al acestora, cred că ar fi șanse mici să se ajungă la o astfel de situație.

Pentru a conduce la un cadru normal de dezbatere electorală. Vă imaginați cum ar arăta o dezbatere electorală doar în doi, cu candidatul A pro și candidatul B anti, cu A profesionist, echilibrat, decent, cu un discurs argumentat, celălalt ...? Și aceasta într-un posibil context economic dificil, marcat de inflație și măsuri impopulare. Acum introduceți în cadru un al treilea candidat, cu un profil mai apropiat de candidatul A. Cu siguranță temele abordate, poziționările, stilul de argumentare ar arăta cu totul altfel, mult mai aproape de ceea ce ne dorim, și, cu siguranță, mult

mai aproape de un context care să ofere posibilitatea luării unei decizii informate, raționale și cumpătate, oricare ar fi aceasta.

O probabilitate mai mare de a primi susținerea partidelor și a aparatelor de partid din teritoriu. Partidelor le va fi foarte greu să accepte să renunțe la un candidat propriu. Totuși, de această dată ar trebui să fie ceva mai ușor decât atunci când procentul obținut la Parlamentare depindea și de calificarea propriului candidat în turul 2. E puțin probabil să fie găsită o variantă ideală de candidat, una care să fie acceptată și susținută de către toate cele patru partide pro-europene. Și, important, și de către actorii majori din societatea civilă. Dacă nu îndrăznesc prea mult, poate și de către „Servicii.” Foarte probabil, șansele ca doi candidați să acopere mai bine aceste preferințe și interese diverse ar fi mai mari.

Pentru a nu spori artificial preferința pentru cealaltă opțiune, cea suveranistă și, pe alocuri, mai puțin democratică și mai extremistă. Într-un context cu doar două opțiuni, foarte probabil, ambele vor fi privite ca având o îndreptățire egală, o parte a electoratului le va percepe a fi la fel de benefice pentru România și români.

La final, o „MICĂ” problemă!

Dacă timp de o jumătate de an partidele au avut dificultăți majore să-și găsească propriul candidat „perfect,” oare îl pot găsi împreună în mai puțin timp? Și dacă au găsit unul, de unde îl mai pot scoate și pe al doilea?

Deocamdată, în baza experiențelor anterioare observate în cazul alegerilor prezidențiale și a celor pentru funcția de Primar General al Municipiului București, șansele par să fie mai degrabă mici. Putem însă spera că „dușul rece” de care partidele tocmai au avut parte va contribui la o schimbare pozitivă.

Poate e nevoie ca societatea, organizațiile neguvernamentale, economice, profesionale, syndicatele să le ajute în acest sens, chiar dacă e posibil ca acest demers să fie contrar intereselor de partid. Și, de ce nu, poate ar fi de preferat ca propunerile să vină

tocmai din aceste zone, iar partidele doar să exprime susținerea, fie aceasta și una coordonată, din interes, pentru candidații propuși.

Calitatea estimărilor sondajelor: Prezidențiale 2024, T1

Data: 2024-12-15

Cuvinte cheie: alegerile-prezidentiale-2024, performanta-sondajelor

Cu ocazia alegerilor din vară am realizat două analize cu privire la calitatea estimărilor produse de sondaje (vezi [aici](#) și [aici](#)). Cred că astfel de analize sunt importante din mai multe puncte de vedere și pentru diferite tipuri de actori (publicul larg, institutele de sondaje, beneficiarii actuali și potențiali ai acestora). Dat fiind faptul că realitatea măsurată de sondaje este adesea fluidă, e important să distingem între sondajele realizate cu mai multe săptămâni înainte de alegeri, cele realizate în săptămâna anterioară alegerilor și cele realizate în ziua alegerilor. Suplimentar, aceste alegeri au făcut notă discordantă ca urmare a unor intervenții majore exercitate de factori externi, cu precădere în ultima săptămână, zilele de vineri-duminică îndeosebi. Prin urmare, mă aștept ca diferențele dintre estimări și rezultatele oficiale să fie mai mari decât de obicei, inclusiv în cazul sondajelor din ziua alegerilor, chiar dacă diferențele vor scădea pe măsură ce ne apropiem de acea zi.

Pentru a ține cont de aceste aspecte, am separat analiza în trei secțiuni: sondaje realizate în primele trei săptămâni de campanie, sondaje realizate în ultima săptămână de campanie și sondaje realizate în ziua votului. Suplimentar, am adăugat o analiză a estimărilor produse prin agregarea datelor de sondaj, respectiv prin modelarea acestor estimări ([modelul realizat de mine](#) și publicat în ziua de vineri).

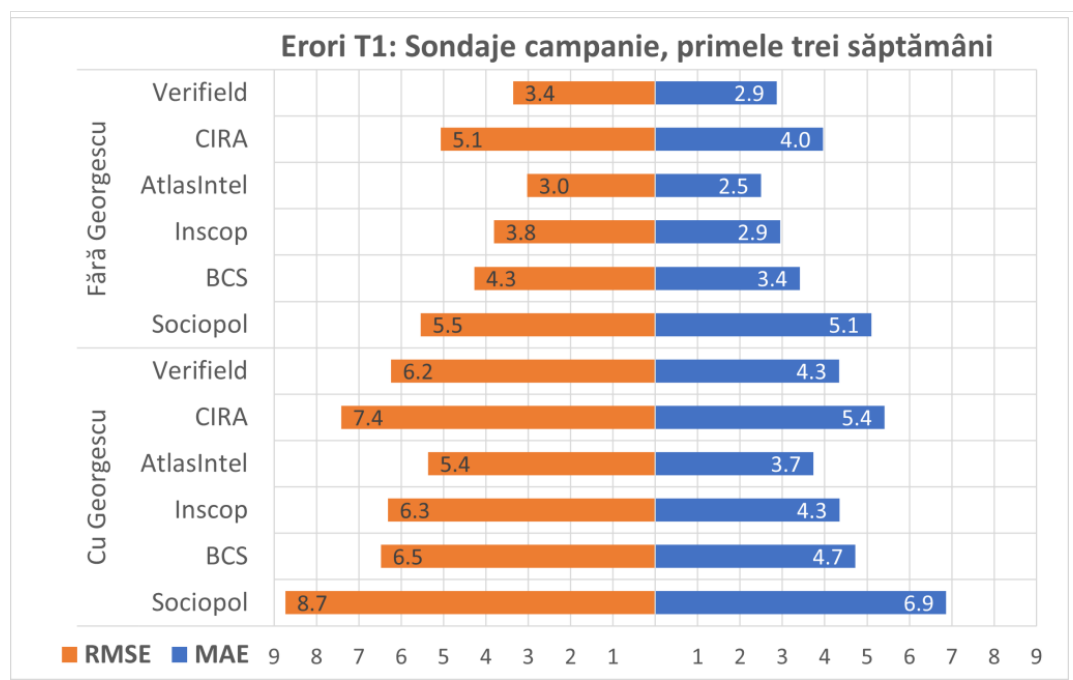
Am considerat doar candidații care au obținut un procent semnificativ de voturi: Ciolacu, Lasconi, Simion, Ciucă, Georgescu, Geoană, Diaconescu, Kelemen și Altul (restul candidaților împreună). Dată fiind situația neobișnuită întâlnită în cazul acestor alegeri, am calculat diferențele în două variante: cu și fără candidatul Georgescu. Desigur, excluderea lui Georgescu elimină doar o parte din efectul produs de evoluția ascendentă rapidă a votului pro Georgescu.

Pentru a măsura calitatea estimărilor am folosit două măsuri: MAE și RMSE. MAE (Mean Average Error) se referă la media erorilor absolute, eroarea fiind definită ca diferența absolută dintre procentul estimat de sondaj și procentul obținut conform BEC. O astfel de măsură indică distanța medie dintre estimările produse de sondaje și rezultatele finale. RMSE (Root Mean Squared Error), rădăcina medie pătratică a erorilor, se calculează ca radical de ordinul doi din suma erorilor ridicate la pătrat. În cazul ambelor măsuri, o valoare mai mare indică o diferență (~eroare) medie mai mare între estimări și rezultatul oficial. Doar diferențele mai mari de trei puncte procentuale pot fi considerate problematice (două puncte procentuale în cazul sondajelor realizate pe eșantioane mari - minimum 2000 „cazuri efective” = numărul cazurilor recalculat pentru situația în care eșantionul a fost de tip SRS (Simple Random Sample), adică simplu aleator.

Estimările sondajelor realizate în primele trei săptămâni de campanie

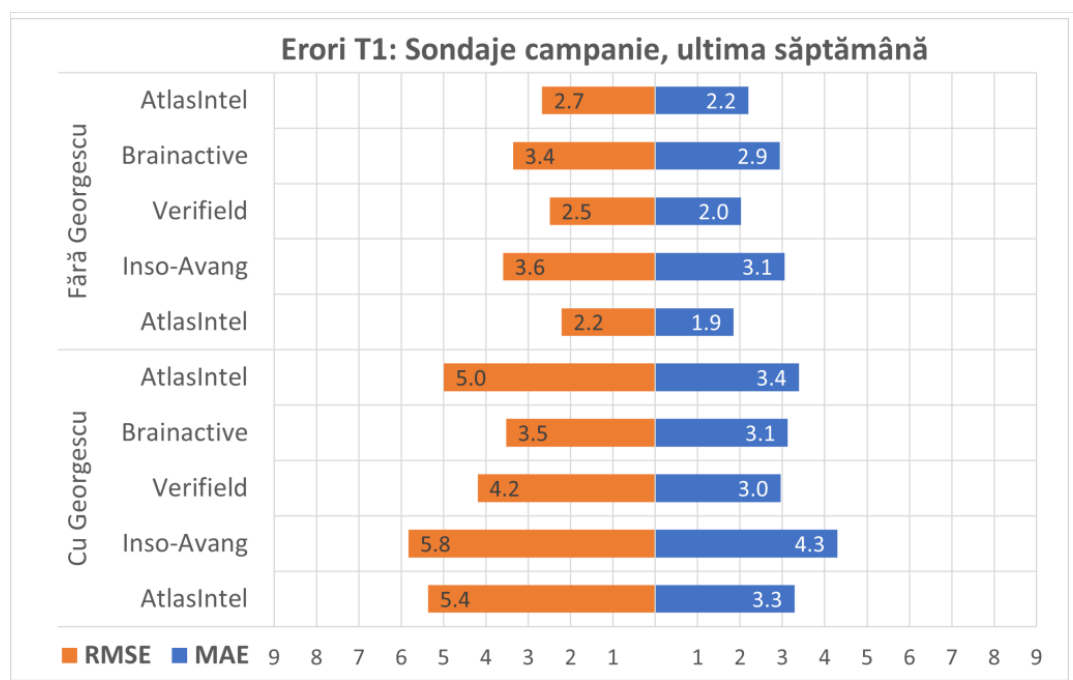
Sondajele realizate în primele trei săptămâni de campanie au produs estimări destul de diferite față de rezultatele oficiale. În acest caz nu e vorba de erori, ci mai degrabă de diferențe față de rezultatele finale. Dat fiind faptul că realitatea s-a schimbat foarte mult pe parcursul campaniei, era de așteptat ca sondajele realizate în urmă cu câteva săptămâni să producă estimări diferite față de rezultatele oficiale. În varianta fără Georgescu, diferențele scad mult și se apropie, mai mult sau mai puțin, funcție de institut, de o diferență medie de trei puncte procentuale.

Important, estimările institutelor sunt influențate și de numărul de zile până la alegeri – sondajele au fost realizate în perioade diferite în interiorul celor trei săptămâni. De exemplu, ultima zi de teren în cazul Sociopol a fost 8 noiembrie, iar în cazul Verifield 14 noiembrie.



Estimările sondajelor realizate în ultima săptămână de campanie

Sondajele realizate în ultima săptămână au fost mult mai precise în medie. În varianta care îl include și pe Georgescu, diferențele medii depășesc uneori pragul de trei puncte procentuale, ajungând până la 4-5 puncte procentuale. În varianta fără Georgescu, diferența medie este în cele mai multe situații mai mică de trei puncte procentuale.

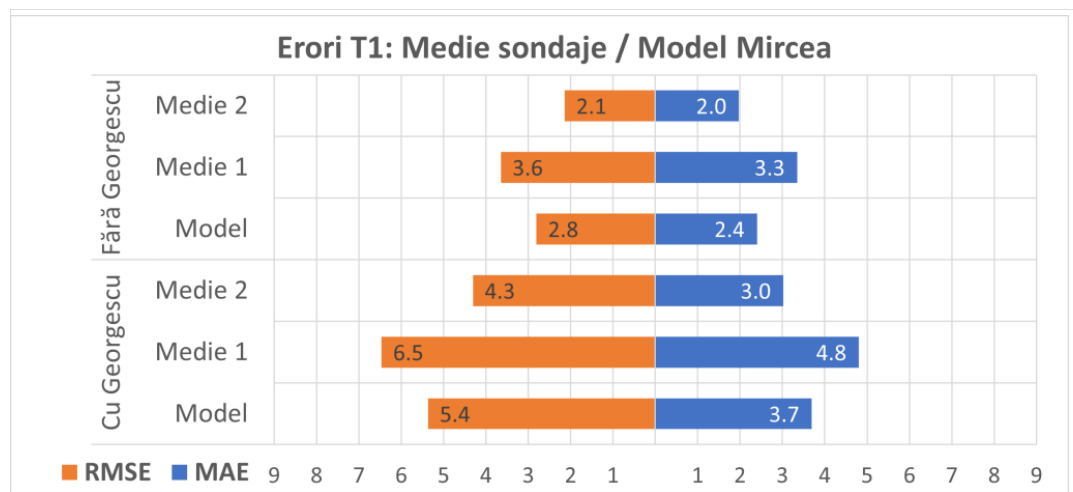


Important: sondajul Brainactive a fost publicat a doua zi după alegeri.

Medii și modelări ale estimărilor din sondaje

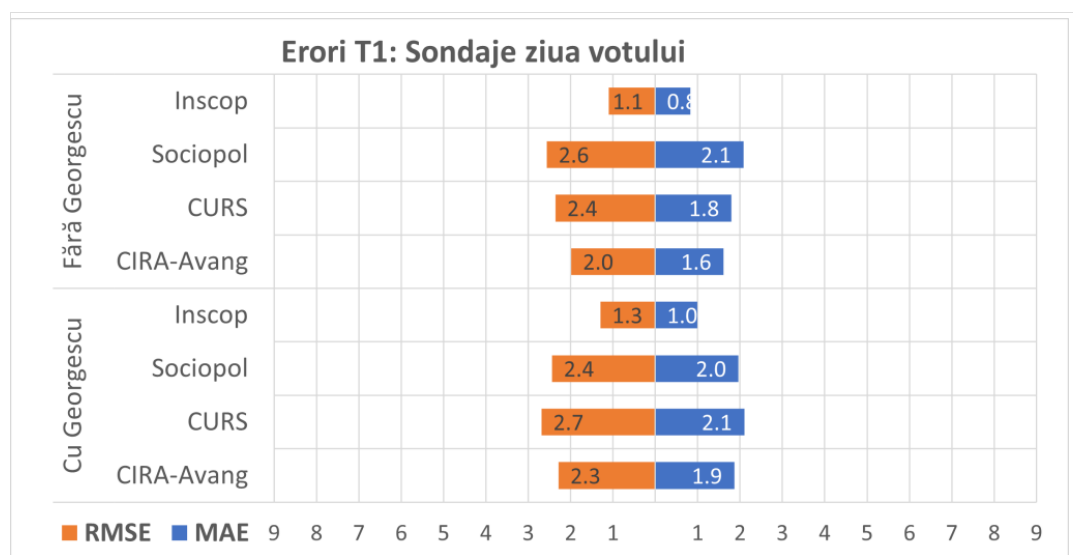
Ce se întâmplă dacă facem o medie a estimărilor din sondaje, respectiv dacă realizăm un model care ia în calcul, alături de alte informații, estimările din sondaje? Și în acest caz am distins între sondajele din săptămâna de campanie și cele mai vechi. Observăm că o medie a sondajelor din primele trei săptămâni de campanie (Medie 1) produce diferențe destul de mari, dar este de preferat sondajelor individuale. O medie a sondajelor din săptămâna anterioară alegerilor (Medie 2) a produs o estimare medie destul de bună, mai ales atunci când Georgescu a fost exclus din analiză. Modelul nu a performat la fel de bine, dar a produs o diferență medie mai mică de trei puncte procentuale. Prin urmare, dacă dorim să știm care sunt intențiile de vot, cel mai bine e să facem o medie a sondajelor din ultima săptămână de campanie. Dacă există, un model, bazat pe sondaje plus alte informații, poate fi util. Dacă procedăm astfel, riscul mediu de a greși și, mai ales,

riscul de a greși în cazul anumitor candidați, e mai mic comparativ cu situația în care ne raportăm la un singur sondaj.



Estimările sondajelor din ziua votului

Sondajele din ziua votului au produs, evident, cele mai bune estimări. Da, au fost erori mari în cazul candidatului Georgescu, dar, media erorilor a fost în toate cazurile mai mică de trei puncte procentuale. Din nou, în varianta fără Georgescu, erorile au fost, în general, ceva mai mici (doar în cazul Sociopol eroarea medie a fost mai mare atunci când l-am exclus pe Georgescu). Cea mai mică eroare medie apare în cazul Inscop, cu mențiune că estimările realizate de acest institut au fost făcute publice câteva ore mai târziu față de celelalte (nu exclud posibilitatea ca unii jurnaliști să le fi avut și/sau să le fi prezentat public ceva mai devreme).



Important: sondajul Inscop a fost publicat pe FB la patru ore după închiderea urnelor; toate celelalte sondaje au fost publicate în intervalul orar 21–22.

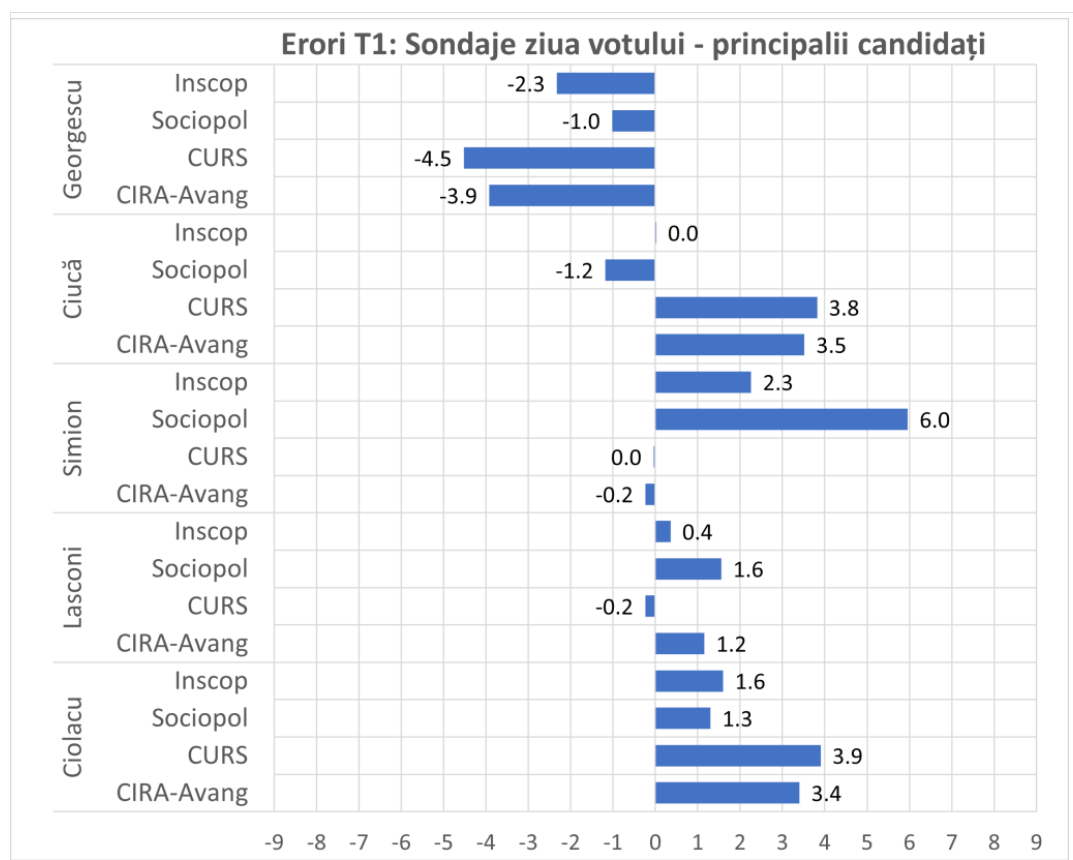
Sub/Supra-estimări în ziua votului

În graficele anterioare, distanța dintre estimări și realitate este calculată ca medie a unor diferențe absolute. Pentru o imagine completă, e necesar să ne uităm și la diferențele punctuale observate în cazul candidaților majori (cei care au obținut mai multe procente). Procedând astfel, putem vedea și sensul diferențelor, respectiv putem face unele inferențe cu privire la sursa erorilor. Astfel, putem observa dacă un candidat a fost sub-estimat sau supra-estimat, dacă un eventual efect de acest tip variază sau nu în funcție de institut și/sau tehnica de sondaj. Pentru a nu lungi prea mult analiza, prezint doar datele cu privire la sondajele realizate în ziua votului.

Observăm că, indiferent de institut, Georgescu a fost sub-estimat, iar Ciolacu a fost supra-estimat. Suplimentar, erorile au fost mai mari în cazul sondajelor realizate la ieșirea de la urne (ceea ce numim exit-poll). E puțin probabil ca eroarea sondajelor de tip

exit-poll să fie cauzată de eroarea de eşantionare (a se vedea [postarea](#) lui Cătălin Stoica de pe FB care arată că eşantionul de secții utilizat de CURS a reprodus foarte bine rezultatul final; în cazul CIRA-Avangarde nu ştim dacă eşantionul de secții de vot a fost la fel de bun). Pe baza acestor observații putem deduce că, în cazul celor doi candidați, eroarea totală de estimare a avut o cauză generală. Foarte probabil e vorba despre ascunderea opțiunii de vot pe diferite căi: (1) rata diferită de acceptare a participării la sondaj, (2) rata diferită de declarare a opțiunii de vot și (3) unii votanți au declarat altă opțiune decât cea reală. Suplimentar, foarte probabil a existat și un efect de metodă (CATI vs. față-în-față la urne) și/sau institut (doar pe baza datelor disponibile, nu putem distinge între aceste efecte).

Procentul obținut de Lasconi a fost estimat corect sau cu mici erori de supra-estimare (Sociopol). Procentul obținut de Ciucă a fost supra-estimată de sondajele la urne (CURS, CIRA-Avangarde), deci, și în acest caz a existat un factor care a acționat sistematic și a distorsionat estimările, similar cu ceea ce am observat în cazul candidatului Ciolacu. Mă refer la efectul indus de combinația dintre metodologie (sondaj față-în-față la urne) și ascunderea opțiunii de vot. Sondajele telefonice (CATI) au supra-estimată opțiunea de vot pentru Simion (nu și cele la urne). În cazul sondajului Sociopol, aproape sigur a existat și o altă sursă de eroare decât cea produsă de tehnica de colectare a datelor (doar așa ar putea fi explicată diferența de șase puncte procentuale dintre estimare și rezultatul oficial în cazul lui Simion).



Important: sondajul Inscop a fost publicat pe FB la patru ore după închiderea urnelor; toate celelalte sondaje au fost publicate în intervalul orar 21-22.

Despre transparența sondajelor

Un comentariu final, aparent marginal, dar extrem de important, cu privire la **transparența sondajelor din România**. Transparența este un criteriu important în funcție de care putem evalua calitatea unui sondaj și profesionalismul unui institut de sondaje. Prin transparență mă refer la tipul de informații care sunt făcute publice, alături de selecția de grafice, și la gradul de detaliere al acestor informații. În cazul sondajelor politice din România, cel mai adesea, sunt prezentate doar informații minimale, fără prea

multe detalii și fără să fie acoperite o serie de informații care sunt foarte importante pentru cineva care dorește să analizeze calitatea datelor produse. Puține institute publică un raport pe propriul site, deci analizele pot fi făcute doar pe baza informațiilor care apar în media. În cazul sondajelor din ziua votului, lucrurile stau chiar și mai rău. Cu referire la ultimele alegeri, nu am găsit niciun raport pe site-ul unui institut, prin urmare am fost nevoit să mă bazez doar pe ceea ce a apărut în media (sau pe FB). O astfel de practică e problematică, cel puțin pentru faptul că, uneori, datele prezentate diferă de la o sursă media la alta. Problema nu e deloc una nouă. A fost semnalată în urmă cu zece ani de Bogdan Voicu într-o serie de [postări](#). S-a schimbat ceva între timp? Sigur da. Din păcate schimbarea a fost în rău. Dacă în urmă cu zece ani mai apărea pe ici pe colo o descriere sumară a metodologiei sondajelor din ziua votului, în prezent nu mai avem nici măcar atât.

Calitatea estimărilor sondajelor: Parlamentare 2024

Data: 2024-12-16

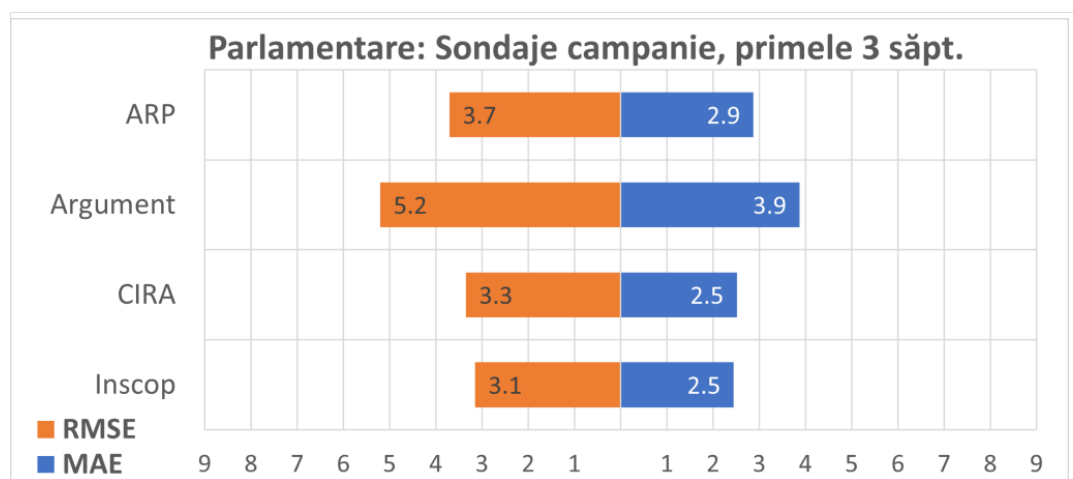
Cuvinte cheie: alegerile-parlamentare-2024, performanta-sondajelor

Continui analiza cu privire la calitatea estimării sondajelor, de această dată cu referire la alegerile parlamentare. Cum acestea produc ceva mai puțină vâlvă în media, deci atrag o audiență mai mică, numărul sondajelor este mult mai mic comparativ cu alegerile prezidențiale. La această „secetă” a contribuit oarecum și ceea ce s-a întâmplat la turul 1 al alegerilor prezidențiale (ascensiunea rapidă a lui Călin Georgescu).

Chiar dacă în cazul alegerilor parlamentare intențiile de vot sunt mai stabile, e important să distingem între sondaje în funcție de perioada în care au fost realizate. Folosim aceeași diviziune temporală: primele trei săptămâni de campanie, ultima săptămână de campanie și ziua votului. Suplimentar, am adăugat o analiză a estimărilor produse prin agregarea datelor de sondaj, respectiv prin modelarea acestor estimări ([modelul realizat de mine](#) și publicat în ziua de vineri). Pentru a măsura diferențele / erorile, am folosit aceleași măsuri (MAE și RMSE). Cu cât aceste măsuri iau o valoare mai mare, cu atât lucrurile stau mai prost din perspectiva calității. Diferențele medii mai mari de trei puncte procentuale pot fi privite ca problematice.

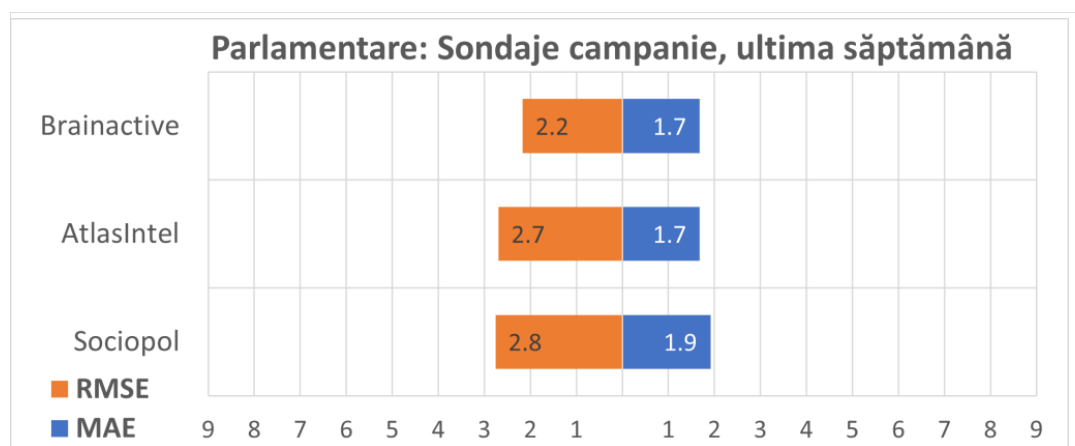
Estimările sondajelor realizate în primele trei săptămâni de campanie

În primele trei săptămâni de campanie au fost publicate doar patru sondaje. Trei dintre ele au produs o calitate medie a estimărilor situată în zona pragului de trei puncte procentuale. În cazul sondajului realizat de Argument, diferența medie a fost ceva mai mare, în principal ca urmare a supra-estimării intenției de vot pentru PSD.



Estimările sondajelor realizate în ultima săptămână de campanie

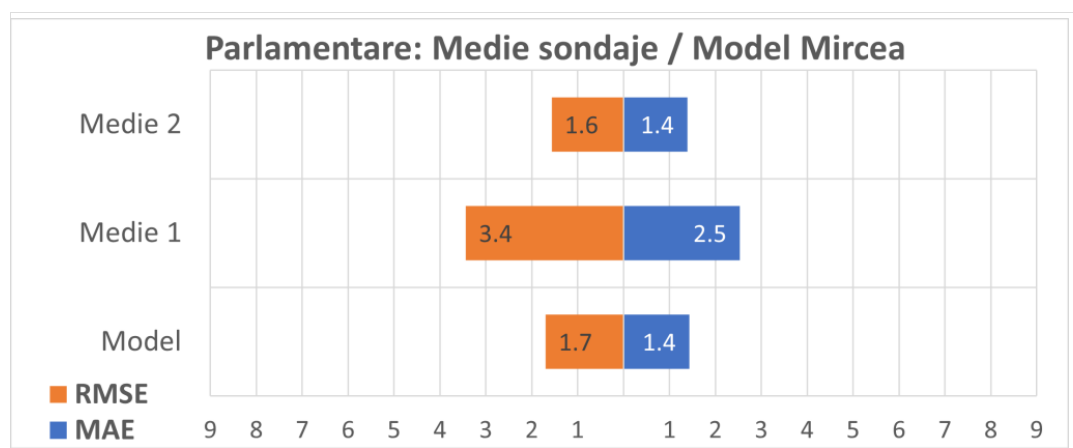
Așa cum mă așteptam, sondajele realizate în ultima săptămână au fost mult mai precise. Toate diferențele medii au fost sub pragul de trei puncte procentuale. Toate institutele au produs, ca medie, estimări de o calitate similară.



Notă: sondajul Brainactive a fost public pe net pentru o perioadă scurtă de timp.

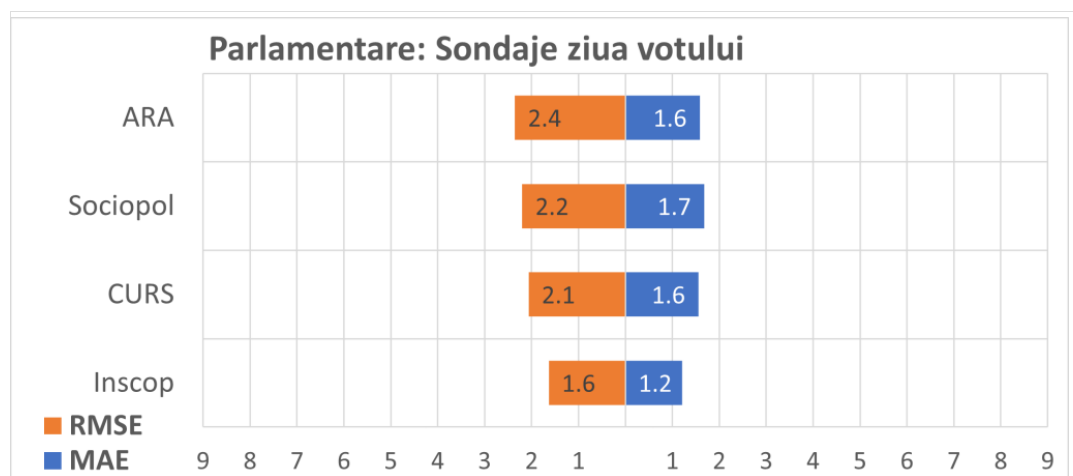
Medii și modelări ale estimărilor din sondaje

Și în cazul alegerilor parlamentare, o medie simplă a estimărilor produse de sondaje conduce la o aproximare mai corectă a intențiilor de vot. O medie a estimărilor din sondajelor realizate în prima parte a campaniei este mai puțin informativă (Medie 1), dar o medie a estimărilor din sondajele realizate în ultima săptămână este foarte informativă (Medie 2). Dat fiind faptul că deja am avut la dispoziție o serie de informații suplimentare (rezultatele oficiale ale alegerilor, matrici de tranziție etc.), modelul construit de mine a produs o estimare medie foarte bună, de o calitate similară cu estimarea medie a sondajelor din ultima săptămână și mai bună decât estimarea medie din sondajele individuale.



Estimările sondajelor din ziua votului

În acord cu așteptările, estimările produse de sondajele realizate în ziua votului au fost mai bune comparativ cu sondajele din săptămâna anterioară. Însă, câștigul de calitate a fost mai degrabă mic. Mai mult, comparativ cu sondajele din ziua votului, o medie simplă a sondajelor din ultima săptămână de campanie, respectiv modelul bazat pe sondaje și alte informații, au condus la estimări mai bune sau la fel de bune (dacă ne raportăm doar la estimările Inscop). Din cele patru sondaje, două au fost telefonice (Sociopol, Inscop) și două la urne (ARA, CURS).



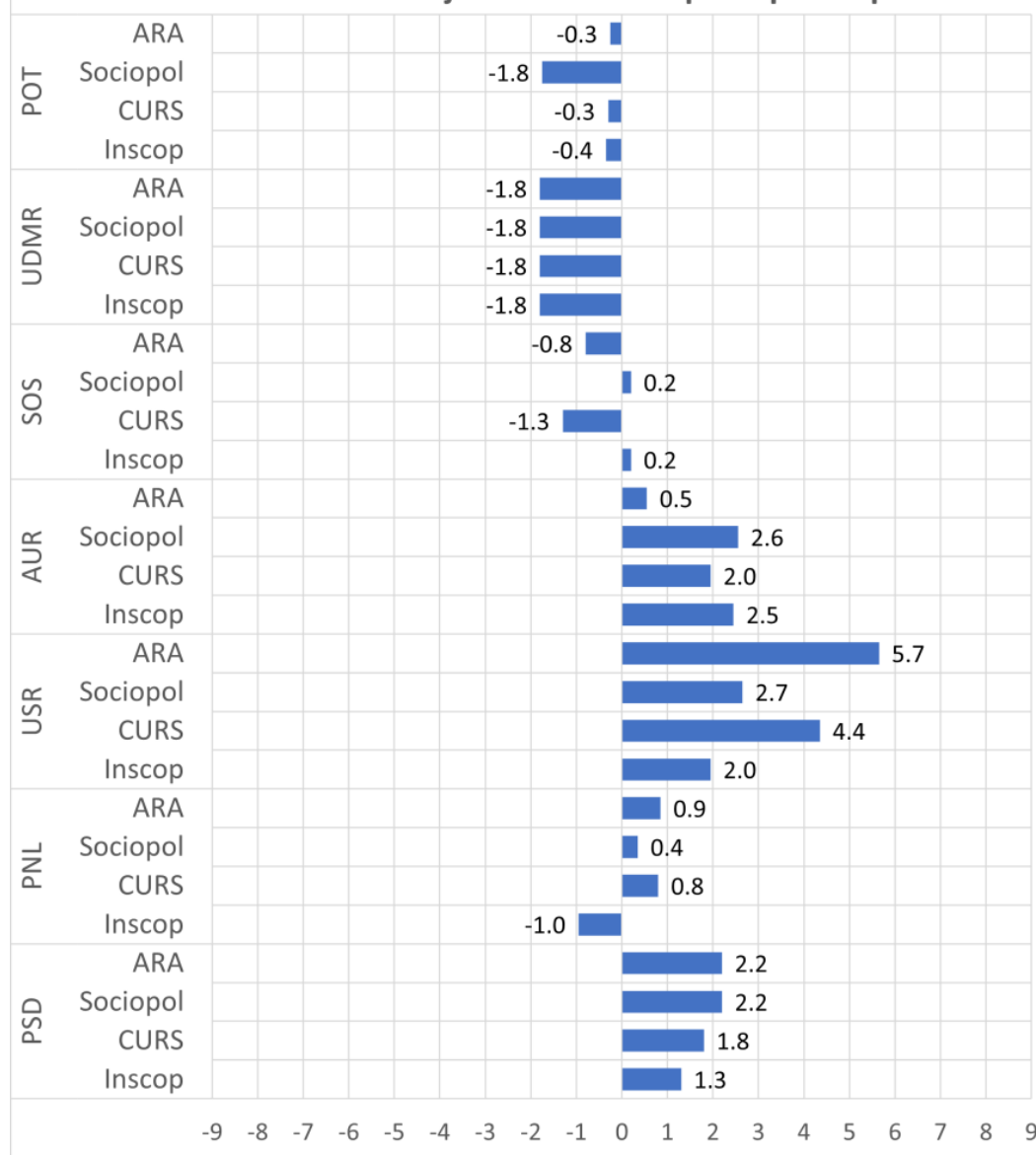
Sub/Supra-estimări în ziua votului

În graficele anterioare, distanța dintre estimări și realitate este calculată ca medie a unor diferențe absolute. Pentru o imagine completă, e necesar să ne uităm și la diferențele punctuale observate. Procedând astfel, putem vedea și sensul diferențelor, respectiv putem face unele inferențe cu privire la sursa erorilor. Astfel, putem observa dacă un partid a fost sub-estimat sau supra-estimat, dacă un eventual efect de acest tip variază sau nu în funcție de institut și/sau tehnica de sondaj. Pentru a nu lungi prea mult analiza, prezint doar datele cu privire la sondajele realizate în ziua votului, alături de câteva concluzii:

- PSD: supra-estimat în toate sondajele cu unul până la două puncte procentuale; foarte probabil cauza acestei erori a fost prezența pe buletinele de vot a unui partid (chiar două) cu nume și siglă similare; o parte a votanților au crezut că au votat PSD, când de fapt au votat alt partid.
- PNL: estimat corect, uneori cu o foarte mică sub-estimare / supra-estimare.
- USR: supra-estimat în toate sondajele cu două până la șase puncte procentuale; foarte probabil a existat un factor general care ține de rata diferită de participare în funcție de educație, vârstă și mediu de rezidență (mai mare în cazul celor mai tineri, educați, din urbanul mare, adică tocmai electoratul USR), respectiv de disponibilitatea mai mare de asumare și declarare a unei preferințe de vot; sondajele la urne au avut erori sensibil mai mari decât cele telefonice; în cazul sondajelor la urne (ARA, CURS), factorii menționați anterior au fost potențați, foarte probabil, de proximitatea fizică, prezența altor persoane, respectiv de distanța socială dintre votanți și operatorii de anchetă (diferențele de vârstă și educație).

- AUR: cel mai adesea, supra-estimat cu aproximativ două puncte procentuale; sursa acestei erori poate avea legătură cu o eventuală supra-corecție a estimărilor; anticipând un efect de indezirabilitate socială, institutule au corectat estimarea de vot în cazul AUR mai mult decât era nevoie; desigur, e posibil și ca o parte a votanților să fi declarat că au votat cu AUR când de fapt au votat cu SOS sau POT (și de aici sub-estimările punctuale în cazul acestor partide).
- SOS: sub-estimare ușoară, doar în cazul sondajelor de la urne, ca urmare e unui efect de indezirabilitate socială.
- UDMR: sub-estimare pe linie, identică, ca și cum estimările de vot pentru UDMR au fost „puse atent cu mâna”; din zece estimări (Senat și CD, estimările de la orele 19-20 și cele finale), nouă au indicat un procent de fix 5% (o estimare a fost 4,8%).
- POT: estimare foarte bună, cu o situație de sub-estimare.

Ereri T1: Sondaje ziua votului - principalele partide



Who does (not) vote in Romania?

Data: 2024-12-21

Cuvinte cheie: rata-de-prezenta-la-vot

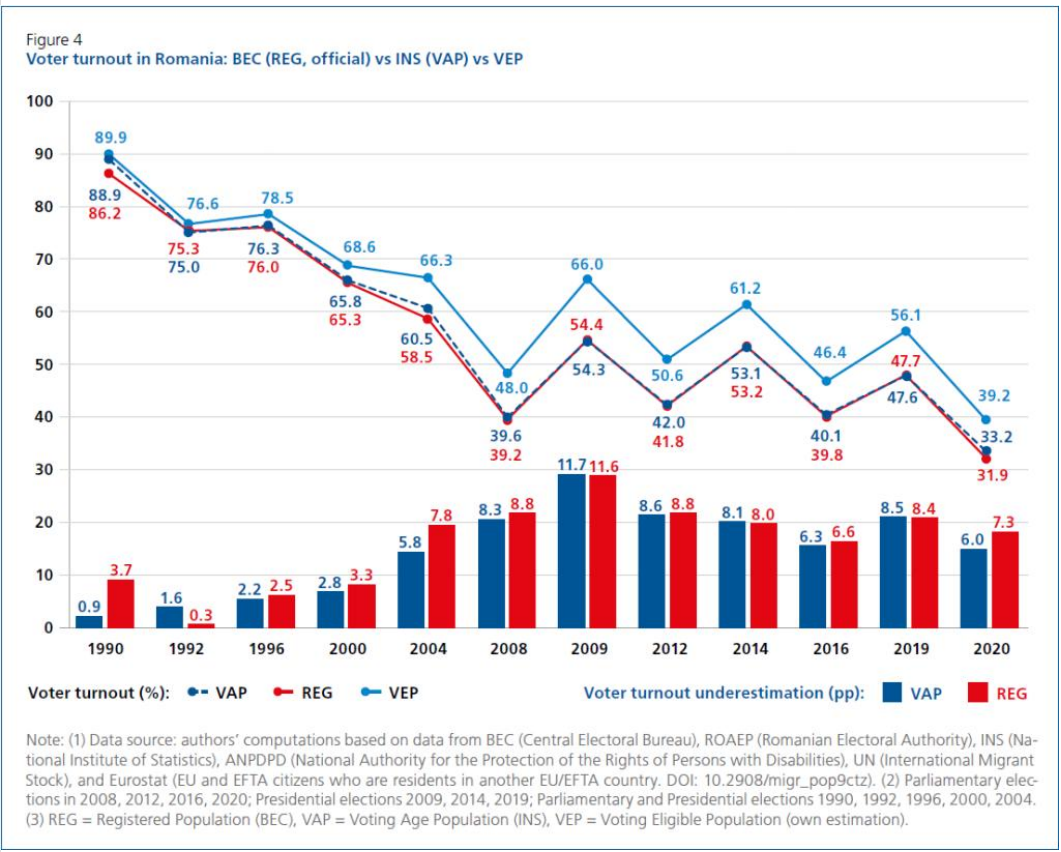
Cu ceva vreme în urmă, împreună cu Cludiu Tufiş, am realizat pentru [FES](#) o analiză a ratei de prezență la vot în cazul alegerilor naționale din România. Raportul a fost publicat luna trecută și poate fi găsit [aici](#). În urma analizei literaturii de specialitate și a datelor disponibile, relativ limitate, am formulat câteva concluzii. O parte a acestora sunt bazate mai mult pe rezultatele unor studii din alte țări decât pe date care măsoară strict situația din România. Pun mai jos trei fragmente din raport alături de un grafic care prezintă dinamica ratei de prezență la vot în cazul alegerilor generale din România.

This report examines, comparatively, the evolution of voter turnout in Romania since the 1990 elections, focusing on the significant decline in electoral participation and on the most important inequalities identified in data from official sources, public opinion surveys, and exit-polls.

Although most post-communist countries have registered voter turnout decline since the founding elections, the decline was the highest in Romania. This can be explained, among others, by high emigration and low levels of interest in politics, linked to low levels of trust in political parties.

The report suggests different solutions to increase voter turnout: remove institutional barriers and improve procedures for diaspora voters, convince political parties to come up with a better

offer for voters, and decrease the voting age to 16 for the European Parliament elections, at least.



O estimare a transferului de voturi dinspre partide spre candidați

Data: 2025-01-08

Cuvinte cheie: alegerile-parlamentare-2024, alegerile-prezidentiale-2024, transferul-de-voturi

Context: După alegerile generale au apărut în media o serie de discuții și informații cu privire la strategiile folosite de partide cu scopul de a crește șansele propriilor candidați de a trece în turul doi al alegerilor prezidențiale, respectiv de a ajunge președinte. Una dintre aceste strategii, de altfel și [recunoscută public](#) de doi reprezentanți importanți ai PSD, a constatat în transferul de voturi dinspre Ciolacu (PSD) spre Simion (AUR) (în prima parte a zilei de vot), posibil și spre Georgescu (în partea a doua a zilei de vot).

Întrebări: Ipoteza transferului de voturi dinspre PSD spre Simion și / sau Georgescu este susținută de datele existente? Dacă da, cât de mare a fost acest transfer? A variat intensitatea transferului în funcție de regiune și mediul de rezidență? Ce alte transferuri importante de voturi au existat?

Metodologie: În extrem de puține situații institutele de sondaje din România fac publice bazele de date. Rapoartele publicate de aceleași institute nu prezintă aproape niciodată tabelul de asociere dintre votul pentru partide și votul pentru candidați. Chiar dacă aș fi avut acces la astfel de date, le-aș fi privit cu anumite rezerve, pentru motive ce țin de limite metodologice și practice. În principal este vorba despre numărul foarte redus de cazuri asociat fiecărei perechi partid-candidat. De exemplu, în cazul unui eșantion de 1000 respondenți, rămânem cu

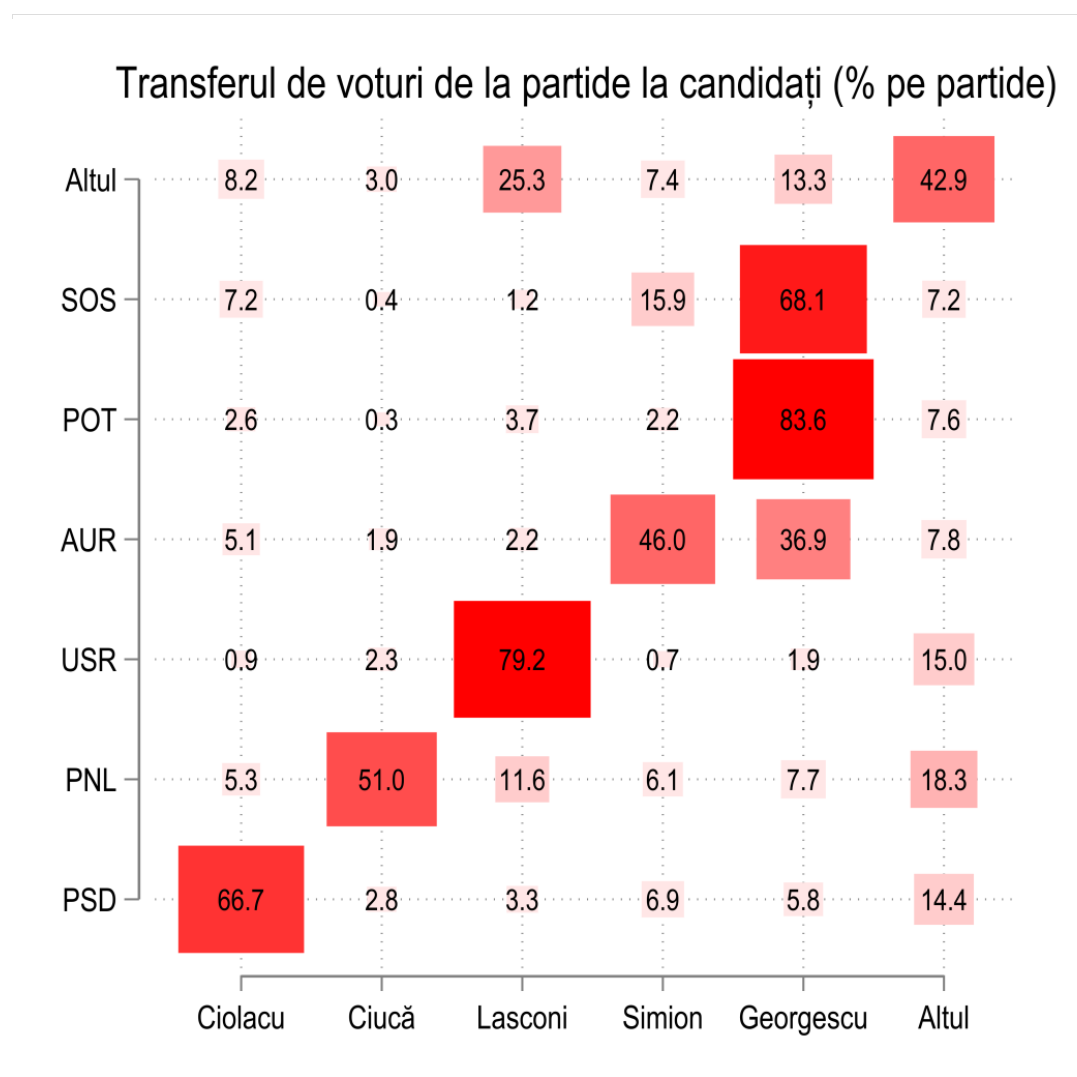
aproximativ 900 alegători care declară că vor vota / au votat, respectiv cu aproximativ 700 de alegători care declară intenția de a vota / votul efectiv pentru un anumit partid. Pentru situația în care avem cinci partide și cinci candidați, rezultă un număr de 25 perechi (partid - candidat), deci o medie de 28 cazuri pe pereche. O estimare care se bazează pe un număr atât de mic de cazuri are o imprecizie foarte mare. Mai mult, în cazul unor perechi „sensibile,” tentația respondenților de a ascunde răspunsul sau de a declara un alt răspuns decât cel real este clar mai mare (de aici și erori de estimare mai mari). Simplu spus, unii alegători nu vor recunoaște că au votat / intenționează să voteze pentru un alt candidat decât candidatul oficial al partidului. În concluzie, o analiză la nivel individual, bazată pe declarații din sondaje, nu este cea mai bună strategie.

O soluție alternativă este analiza / inferența ecologică. E vorba de o analiză realizată pe date la nivel agregat cu scopul de a produce inferențe cu privire la situația de la nivel individual. În domeniul studiilor electorale analiza este cunoscută sub numele de „voter transition matrix / probabilities.” Pentru a reduce cât mai mult posibilitatea realizării unor inferențe incorecte, e important să folosim date situate la un nivel de agregare cât mai redus. Prin urmare, toate analizele asociate acestui text sunt realizate pe datele oficiale la nivel de secție de vot (turul 1 de la Prezidențiale, respectiv Camera Deputaților). Am considerat toate secțiile de vot, inclusiv cele din străinătate. Pentru a simplifica prezentarea datelor, am redus analiza la partidele / candidații care au obținut un număr relativ mai mare de voturi.

Două precizări importante: (1) Analiza nu permite identificarea cauzei transferului de voturi: un alegător poate vota cu un alt candidat decât cel propus de partid în baza unor motivații și preferințe personale sau, dimpotrivă, pentru că așa i-au cerut reprezentanții partidului preferat. (2) Pentru că ne interesează transferul de voturi dinspre partide spre candidați, am realizat analiza ca și cum alegerile parlamentare ar fi avut loc înainte de cele prezidențiale.

Rezultate: Valorile din grafic sunt procente pe linii (totalul pe fiecare linie e 100%), deci indică, pentru fiecare partid, ponderea

voturilor care s-au dus spre un anumit candidat. La nivelul fiecărei perechi (celule / combinații partid + candidat), mărimea relativă a transferului este comunicată vizual prin intermediul intensității culorii și a mărimii pătratului din fundal. O culoare roșie mai intensă, respectiv un pătrat mai mare, indică o pondere mai mare a voturilor transferate dinspre un partid înspre un candidat.



Cel mai mare transfer de voturi apare în cazul POT: 84% dintre voturile primite de POT au mers la Georgescu. Pe locurile următoarea se situează perechile USR – Lasconi (aceasta primește

79% din voturile obținute de USR), SOS - Georgescu (primește 68% din voturile obținute de SOS), PSD - Ciolacu (67%), PNL - Ciucă (51%), AUR - Simion (46%) și AUR - Georgescu (37%).

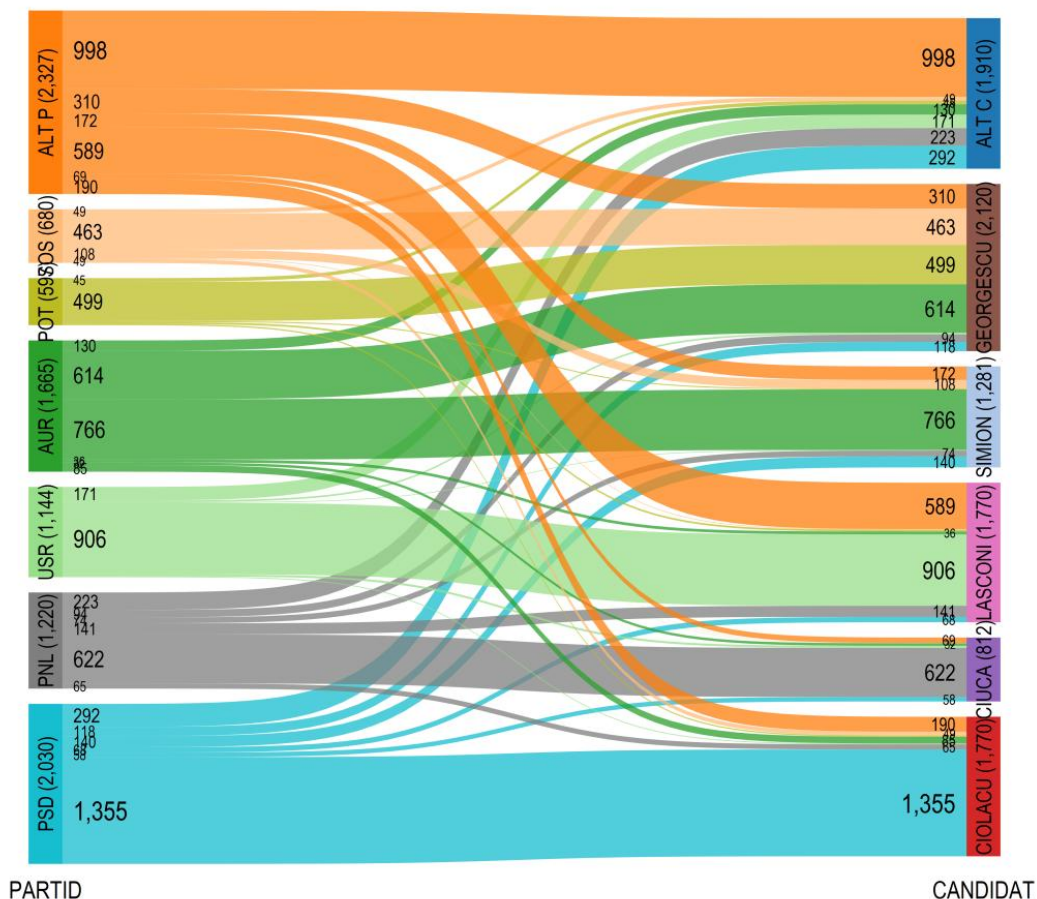
Important pentru întrebările de interes aici, 7% dintre voturile PSD merg la Simion, iar 6% la Georgescu. În cazul lui Simion procentul este mai mare (9-10%) în regiunea Vest și în ruralul din regiunea Centru. În cazul lui Georgescu procentul este mai mare (9-10%) în regiunea Nord-Vest și în urbanul din regiunea Sud-Est.

Alte transferuri importante: PNL - Lasconi (12%), SOS - Simion (16%), Alt partid - Lasconi (25%), Alt partid - Georgescu (13%), PSD - Alt candidat (14%), PNL - Alt candidat (18%), USR - Alt candidat (15%), Alt partid - Alt candidat (43%).

Pun mai jos un grafic alternativ de ilustrare a transferului de voturi ca valori absolute. Totalurile nu reproduc exact valorile oficiale deoarece graficul include la categoriile Alt partid și Alt candidat, pe lângă voturile obținute de restul partidelor și candidaților, și intrările, respectiv ieșirile de voturi (cei care au votat doar la una dintre cele două alegeri).

Un exemplu de citire a rezultatelor în cazul PSD: din cele 2030 mii de voturi primite de PSD, 1355 mii au mers spre Ciolacu, 140 mii spre Simion, 118 mii spre Georgescu și 292 mii spre alți candidați (Geoană în principal). Și un exemplu în cazul voturilor obținute de Georgescu: din cele 2120 mii voturi obținute de Georgescu, 614 mii au venit de la votanții AUR, 499 mii de la POT, 463 mii de la SOS, 118 mii de la PSD, 94 mii de la PNL și 310 mii de la alt partid (sau de la cei care au votat doar la turul 1).

Transferul de voturi de la partide la candidați (mii)



La sugestia lui Mircea Kivu, adaug mai jos graficul cu estimarea transferurilor (sau conversiilor) în cazul secțiilor de vot din străinătate. Pentru mai multe clarificări, se pot vedea și comentariile de pe blog.

Diaspora										
	GEORGESCU	LASCONI	CIOLACU	SIMION	CIUCA	KELEMEN	GEOANA	ALT_C	NET_EXITS	
PSD	6.8	8.4	37.8	3.0	29.2	0.1	3.1	5.4	6.0	30730
AUR	56.7	4.0	1.4	23.4	1.2	0.3	2.5	5.4	5.2	201280
PNL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
USR	7.5	66.9	1.8	2.4	9.4	0.2	4.1	2.4	5.3	185485
POT	89.0	1.4	0.4	1.9	0.1	0.1	0.6	1.2	5.3	104439
SOS	61.2	2.1	1.0	23.4	0.2	0.1	1.4	5.3	5.3	121639
UDMR	20.3	27.3	3.4	2.9	2.5	24.9	4.3	2.4	12.1	11606
ALT_P	7.1	50.0	1.6	4.2	4.3	0.2	17.6	10.1	4.9	108609
NET_ENTRIES	51.1	23.9	2.0	11.4	3.7	0.5	4.4	2.8	0.0	74929
	345925	214033	22893	96339	36934	4738	38966	38066	40823	

„Da’ liberalii un’ se duc când se duc?”

Data: 2025-01-11

Cuvinte cheie: alegerile-parlamentare-2024, diaspora, transferul-de-voturi

Evident, jocul de cuvinte din titlu este inspirat din nuvela Bunicul a lui Delavrancea și are doar scopul de a stârni interesul pentru lectura analizei. Postarea nu are nicio legătură cu țara cocorilor și nici cu trecerea în neființă a cuiva. Discuția este strict despre votanții PNL (secundar FD) din diaspora.

La alegerile parlamentare din 2024, în diaspora a fost aprobată doar una din cele două liste de candidați ai PNL, cea pentru Senat. Lista PNL pentru Camera Deputaților a fost respinsă pentru că nu a inclus nicio femeie. FD a pățit același lucru, în cazul listei pentru Senat. Desigur, o astfel de situație este una nefericită pentru PNL și votanții săi (nu și pentru celelalte partide aflate în competiția electorală). Însă, din punct de vedere științific, datele rezultate în urma acestei situații sunt utile. Fără să-și propună acest lucru, PNL și FD ne-au oferit posibilitatea să analizăm experimental transferul de voturi dinspre PNL / FD spre alte partide. Important, putem face analiza pe baza comportamentelor votanților și nu a declarațiilor (potențialilor) alegători (așa cum sunt ele măsurate prin sondaje). Bonus, putem studia fenomenul de interes pe toată populația de votanți din diaspora, nu doar pe un eșantion selectat din această populație. Pentru că apropierea de o abordare de tip experimental este întotdeauna de dorit, iar în științele sociale, pentru motive evidente, design-urile de tip experimental sunt relativ rare, putem spune din toată inima „bogdaproste.”

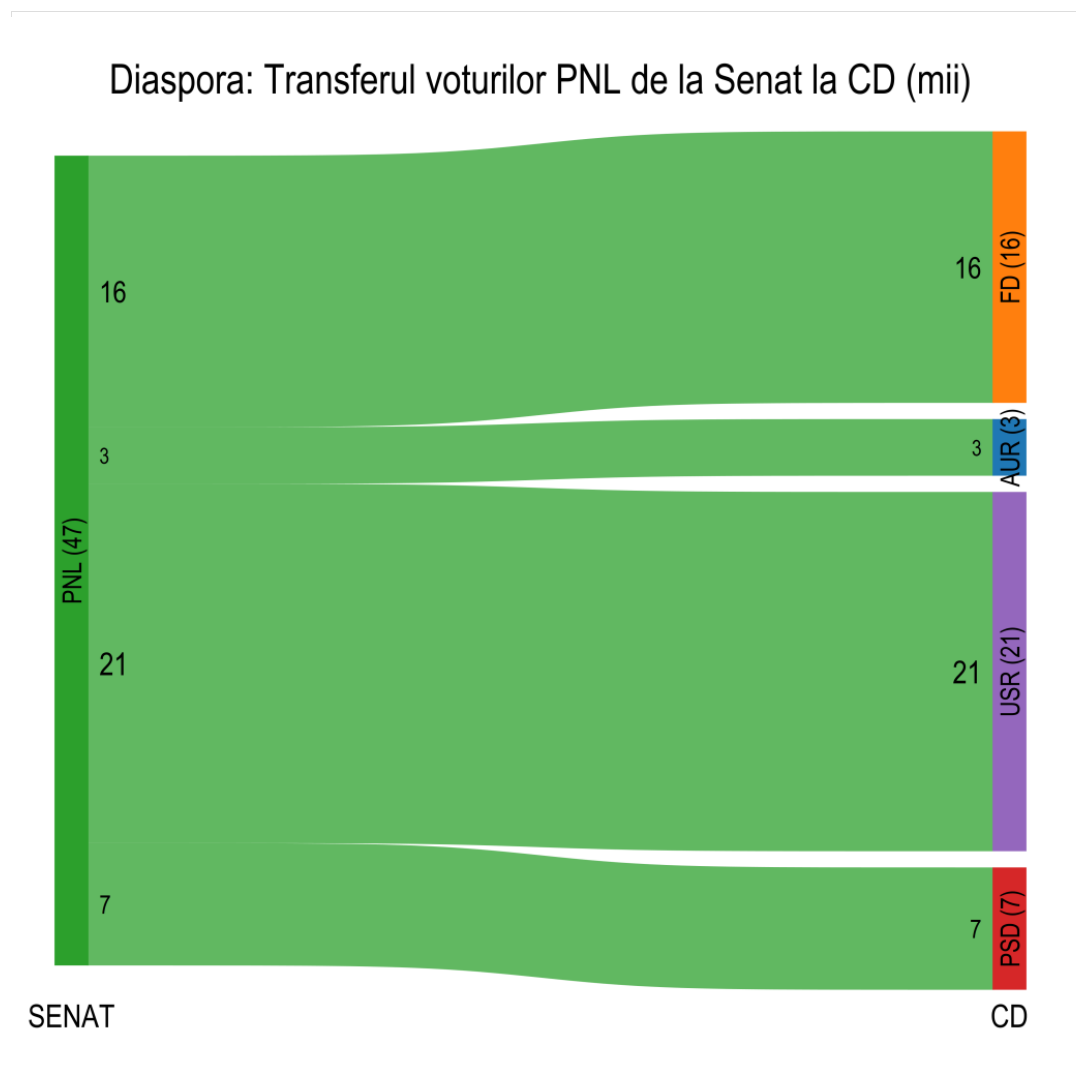
Pentru mai multe detalii metodologice cu privire la tipul de analiză realizat se poate consulta această [postare](#). Valorile prezentate sunt calculate pe baza datelor oficiale la nivel de secție de vot din diaspora. Valorile din tabele sunt procente pe linii (suma e 100% pe fiecare linie). Pentru că aceeași alegători au votat la ambele Camere, ponderea intrărilor și ieșirilor nete de voturi a fost aproape zero, prin urmare am ignorat aceste categorii.

În primul tabel prezint transferurile de voturi dinspre Senat înspre Camera Deputaților (secții de vot din străinătate). Mai exact, e vorba de ratele de transfer sau de conversie (într-o [denumire preferată de Adrian Hatos](#)). Se observă că, în cele mai multe cazuri, peste 90% din cei care votează un partid la Senat votează același partid și la Camera Deputaților. În cazul DREPT și REPER ponderea e puțin mai mică (77% și 85%). Pentru că PNL nu a avut candidați la Camera Deputaților, voturile obținute de acest partid la Senat s-au îndreptat la Camera Deputaților spre USR (39%), FD (26%), PSD (11%) și AUR (7%).

	CD_PSD	CD_PNL	CD_USR	CD_AUR	CD_SOS	CD_POT	CD_REPER	CD_SENS	CD_DREPT	CD_UDMR	CD_FD	CD_ALTUL	CD_VOTNUL	
SE_PSD	92.9	0.0	0.4	0.5	0.3	0.2	0.1	0.1	0.0	0.0	0.2	5.3	0.0	23237
SE_PNL	10.7	0.0	38.8	6.6	2.6	1.9	1.4	1.0	0.6	0.5	25.8	8.5	1.6	57429
SE_USR	0.4	0.0	95.2	0.6	0.4	0.3	0.4	0.3	0.2	0.1	1.1	0.7	0.2	167376
SE_AUR	0.2	0.0	0.2	95.2	0.7	0.8	0.3	0.3	0.5	0.2	0.2	1.0	0.5	197731
SE_SOS	0.2	0.0	0.3	3.9	93.1	0.8	0.1	0.2	0.2	0.1	0.2	0.6	0.3	124262
SE_POT	0.3	0.0	0.4	1.6	1.0	94.9	0.1	0.1	0.3	0.1	0.1	0.8	0.2	104598
SE_REPER	1.8	0.0	8.2	1.2	0.6	0.5	84.7	1.9	0.1	0.1	0.6	0.3	0.1	14372
SE_SENS	1.0	0.0	1.2	1.0	0.6	0.5	0.9	93.8	0.2	0.1	0.1	0.4	0.2	28354
SE_DREPT	2.0	0.0	2.0	2.2	5.0	2.8	0.2	0.6	77.1	0.0	0.6	7.2	0.2	17336
SE_UDMR	1.5	0.0	1.2	0.9	0.5	0.3	0.1	0.1	0.1	94.2	0.7	0.3	0.1	11133
SE_FD	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
SE_ALTUL	0.4	0.0	0.4	0.8	0.1	0.1	0.0	0.0	0.0	0.0	0.1	97.6	0.3	19380
SE_VOTNUL	3.1	0.0	4.4	8.1	2.8	2.7	0.2	1.1	0.2	0.3	0.2	3.8	73.0	5849
	30730	0	185485	201280	121639	104439	14808	29165	15676	11606	17779	31181	7276	

În termeni absoluți, cele 57 mii voturi obținute de PNL în diaspora-Senat s-au împărțit la Camera Deputaților astfel: 7 mii

PSD, 21 mii USR, 3 mii AUR și 16 mii FD (restul de 10 mii s-au împărțit între celelalte partide).



În al doilea tabel prezint transferurile de voturi dinspre Camera Deputaților înspre Senat (secții de vot din străinătate). Se observă că, în multe situații, în jur de 90% dintre cei care votează un partid la Camera Deputaților votează același partid și la Senat. În cazul REPER ponderea e puțin mai mică (80%). În cazul PSD ponderea e chiar și mai mică: doar 54% dintre voturi merg spre PSD la Senat, iar 39% merg spre PNL la Senat. Pentru că FD nu a avut candidați

la Senat, voturile obținute de FD la Camera Deputaților merg la Senat aproape în totalitate la PNL (97%).

	SE_PSD	SE_PNL	SE_USR	SE_AUR	SE_SOS	SE_POT	SE_REPER	SE_SENS	SE_DREPT	SE_UDMR	SE_FD	SE_ALTUL	SE_VOTNUL	
CD_PSD	53.5	38.5	2.8	2.0	1.0	0.9	0.2	0.1	0.2	0.1	0.0	0.4	0.3	30730
CD_PNL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
CD_USR	0.5	7.9	85.4	1.6	1.0	0.8	0.6	0.6	0.4	0.2	0.0	0.8	0.3	185485
CD_AUR	0.3	0.9	0.4	92.6	1.4	1.3	0.3	0.3	0.6	0.1	0.0	1.1	0.6	201280
CD_SOS	0.4	0.5	0.3	1.2	95.9	0.7	0.1	0.1	0.2	0.0	0.0	0.5	0.2	121639
CD_POT	0.3	0.5	0.3	4.3	0.6	93.2	0.1	0.1	0.3	0.0	0.0	0.2	0.1	104439
CD_REPER	3.2	4.5	9.8	1.1	0.6	0.3	79.8	0.2	0.1	0.1	0.0	0.2	0.0	14808
CD_SENS	1.0	2.3	5.2	0.7	0.5	0.4	0.5	88.9	0.1	0.1	0.0	0.2	0.1	29165
CD_DREPT	1.2	1.9	1.1	0.9	0.7	6.0	0.1	0.1	87.9	0.0	0.0	0.0	0.0	15676
CD_UDMR	2.3	4.2	1.6	1.0	0.6	0.6	0.2	0.2	0.2	88.5	0.0	0.3	0.4	11606
CD_FD	0.2	96.5	3.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	17779
CD_ALTUL	8.9	24.8	6.9	3.7	3.1	2.2	1.0	0.5	2.3	0.2	0.0	46.2	0.2	31181
CD_VOTNUL	5.0	14.2	5.5	3.3	8.2	4.1	0.3	1.4	2.3	0.2	0.0	5.3	50.1	7276
	23237	57429	167376	197731	124262	104598	14372	28354	17336	11133	0	19380	5849	

Profile electorale: Prezidențiale 2024, T2

Data: 2025-01-26

Cuvinte cheie: alegerile-prezidentiale-2024, intentii-de-vot, profile-electorate

În ultimele postări am analizat câteva teme pe baza datelor din ancheta IRES de la finalul anului 2024 (vezi [aici](#) și [aici](#); mulțumiri IRES pentru accesul la baza de date și acordul de publicare a acestor analize). Unele dintre temele analizate par să nu aibă o legătură una cu alta. Temele au totuși un numitor comun: există suspiciuni rezonabile că au o legătură cu preferința de vot. În analiza de față investighez tocmai această legătură, de această dată în cazul turului doi al alegerilor prezidențiale. Întrebarea de vot a fost următoarea:

Dacă turul al doilea al alegerilor prezidențiale nu ar fi fost anulat, care ar fi fost opțiunea dumneavoastră?

Pentru această analiză am grupat răspunsurile în trei categorii: ar fi votat cu Lasconi, ar fi votat cu Georgescu și alt răspuns (nu ar fi votat, vot anulat, nu știu, nu răspund). În grafice am prezentat doar valorile asociate cu votul pentru Lasconi / Georgescu, respectiv diferența dintre aceste valori.

Analiza este grupată în trei secțiuni:

- (1) O comparație a profilelor electorale ale candidaților Elena Lasconi și Călin Georgescu.
- (2) O comparație a factorilor care au contribuit la formarea opțiunilor de vot.

(3) O comparație a probabilităților de vot în funcție de principalii determinanți ai votului.

Care au fost intențiile de vot în funcție de fiecare dintre sub-populațiile pe care le-am avut în vedere? Graficul de mai jos ne răspunde la această întrebare. Primele două coloane (Variabila și Categoria) definesc sub-populația de interes. Următoarele două coloane arată intențiile de vot pentru fiecare dintre cei doi candidați, adică vestitele „procente pe linii!” Diferența până la 100% o reprezintă cei care au oferit un alt răspuns. Ultima coloană prezintă diferența dintre ponderea voturilor pentru Georgescu și ponderea voturilor pentru Lasconi.

Graficul permite compararea simultană a preferințelor de vot în funcție de candidat și o sub-populație, respectiv, la nivelul unui candidat, în funcție de o sub-populație. De exemplu, putem observa că 28% dintre femei declară că ar fi votat cu Lasconi și 47% cu Georgescu (restul de 25% dau un alt răspuns). Pentru că aproape aceleași valori apar și în cazul bărbaților (30%, 50%, 20%), putem spune că intențiile de vot ale celor două sub-populații sunt similare. Cu referire la nivelul de educație formală observăm că (1) ponderea voturilor pentru Georgescu scade pe măsură ce crește nivelul de educație, (2) ponderea voturilor pentru Lasconi crește pe măsură ce crește nivelul de educație, (3) dacă ar vota doar cei cu studii elementare, Georgescu ar fi câștigat detașat și (4) dacă ar vota doar cei cu studii superioare, Lasconi ar fi câștigat detașat.

Vă las pe voi să descoperiți alte diferențe importante. Nu uitați însă că datele din acest grafic indică doar posibilitatea existenței unei legături între intenția de vot și caracteristicile analizate. Fie următorul exemplu: Observăm că populația din mediul rural votează într-o măsură mai mare cu Georgescu. Putem spune că acești alegători votează cu Georgescu din cauză că trăiesc în rural? Nu neapărat. Ar trebui să verificăm dacă nu există o altă caracteristică (educația?) care e asociată simultan cu mediul de rezidență și intenția de vot. Mai apar diferențe de vot în funcție de mediul de rezidență atunci când controlăm nivelul de educație (adică analizăm relația dintre vot și mediul de rezidență separat pentru fiecare nivel de educație)? Anticipez, datele spun că nu. Cei

din rural votează într-o măsură mai mare cu Georgescu pentru că nivelul lor de educație e mai scăzut, nu pentru că locuiesc în rural.

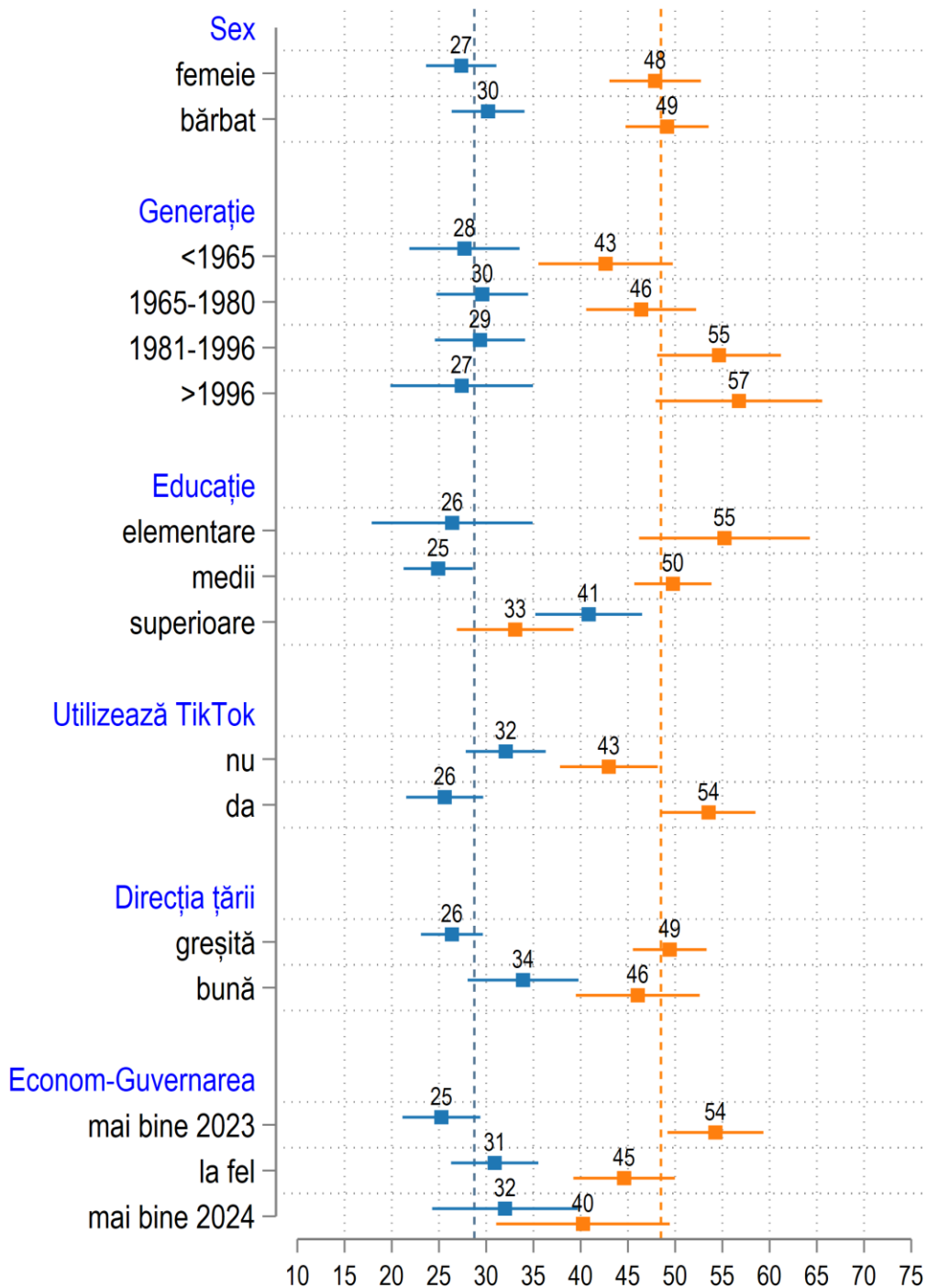
Variabila	Categoria	Lasconi	Georgescu	Georgescu - Lasconi
Sex	femeie	28	47	
	bărbat	30	50	
Generație	<1965	28	40	
	1965-1980	27	50	
	1981-1996	31	54	
	>1996	31	53	
Școlaritate	elementare	17	65	
	medii	24	50	
	superioare	59	19	
Mediu	rural	22	57	
	urban	35	41	
Regiune	București	39	38	
	Intra-Carpatic	31	50	
	Muntenia	20	54	
	Moldova	35	40	
Folosește TikTok	nu	37	36	
	da	21	61	
Direcția țării	greșită	22	54	
	bună	46	33	
2024 vs. 2023: România	mai bine 2023	19	60	
	la fel	37	39	
	mai bine 2024	38	36	
Trai personal	nemulțumit	16	60	
	mulțumit	39	39	
2024 vs. 2023: banii	mai bine 2023	15	63	
	la fel	33	44	
	mai bine 2024	37	39	
Principala îngrijorare (2024)	boală	31	44	
	inflație	27	49	
	criză politică	29	49	
	război	34	43	
	altă frică	22	58	
Reprezentare comunism: deprivare	mică [-1;0.5]	12	57	
	mare (0.5;0.8]	15	61	
	foarte mare (0.8;1]	38	41	
Imaginaea comunismului: lucrurile bune	cel mult neutră [-1;0]	63	17	
	ușor pozitivă (0;0.5]	44	33	
	pozitivă (0.5;0.9]	29	43	
	total pozitivă [1]	14	66	
Era mai bine în 1980-1989 decât în prezent	nu	40	36	
	da	10	70	
1989 = lovitură de Stat	nu	39	38	
	da	22	56	
1989 = eveniment negativ	nu	35	42	
	da	17	61	
Preferă o societate democratică	puțin (0;0.3]	7	72	
	mult (0.3;0.7]	19	59	
	foarte mult (0.7;1]	40	36	
Preferă: Statul stabilește prețurile & salariile	deloc/puțin [-1;-0.3]	32	50	
	oarecum (-0.3;0.3]	26	50	
	mult (0.3;0.7]	35	39	
	foarte mult (0.7;1]	23	56	
Raport salariu maxim / minim	NR	38	36	
	<2	21	55	
	2-2.9	25	52	
	3-4.9	34	47	
	5+	43	36	
Încredere profesii: sistemul public	foarte mică [-1;-0.5]	17	56	
	mică [-0.5;0]	25	52	
	mare [0;0.5]	33	45	
	foarte mare (0.5;1]	35	45	
Încredere profesii: mediul privat	foarte mică [-1;-0.5]	12	61	
	mică [-0.5;0]	23	54	
	mare [0;0.5]	33	44	
	foarte mare (0.5;1]	35	46	
Încredere profesii: guvernantă	foarte mică [-1;-0.5]	22	52	
	mică [-0.5;0]	27	54	
	mare [0;0.5]	39	39	
Încredere profesii: preot	foarte mică [-1;-0.5]	32	44	
	mică [-0.5;0]	33	48	
	mare [0;0.5]	30	44	
	foarte mare (0.5;1]	20	60	

Pe baza analizelor anterioare am selectat caracteristicile în funcție de care intențiile de vot variază semnificativ. Am construit un model de predicție a intențiilor de vot în care am ținut cont simultan de toate aceste caracteristici. Valorile din următoarele trei grafice sunt predicții ale intențiilor de vot produse de acest model. Datele din acest grafic ne permit să formulăm, cu unele limite, afirmații (potențial) cauzale. La modul general, citim datele astfel: „Celelalte caracteristici considerate fiind aceleași, comparativ cu categoria 1 (de exemplu femeie) categoria 2 (bărbat) are o probabilitate mai mare / mai mică / similară de a vota cu X.” Câteva exemple de interpretare: Dacă eliminăm celelalte diferențe considerate în analiză, femeile și bărbații au preferințe de vot similare; preferința pentru Lasconi e mai mare în cazul votanților cu studii superioare; preferința pentru Georgescu scade pe măsură ce crește nivelul de educație formală; raportat doar la cei cu studii superioare, 41% ar vota cu Lasconi și 33% cu Georgescu; cei care utilizează TikTok au șanse mai mari să voteze cu Georgescu.

Intenția de vot pentru Georgescu crește pe măsură ce scade vârsta, schimbările economice și politice din ultimul an sunt evaluate negativ, perioada comunistă este văzută pozitiv, preferința pentru democrație este mai mică, se dorește ca Statul să stabilească prețurile și salariile și există o încredere mai mare în preoți. **Intenția de vot pentru Lasconi** crește pe măsură ce perioada comunistă este văzută negativ (deprivare, puține lucruri bune), preferința pentru democrație este mai mare, există o încredere mai mare în profesiile din sistemul public și mai mică în preoți.

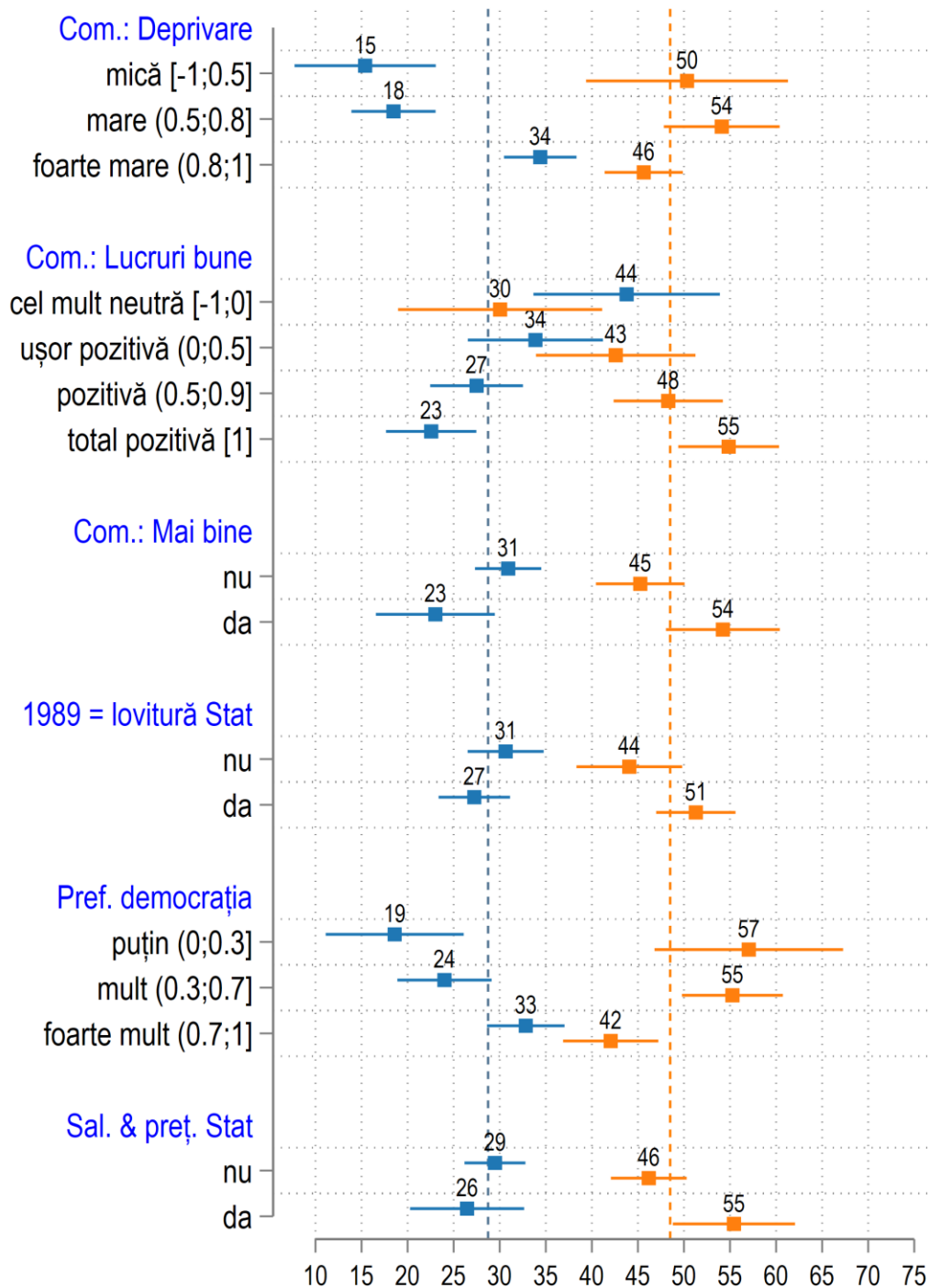
Cu cine ar fi votat la turul 2

■ Lasconi ■ Georgescu

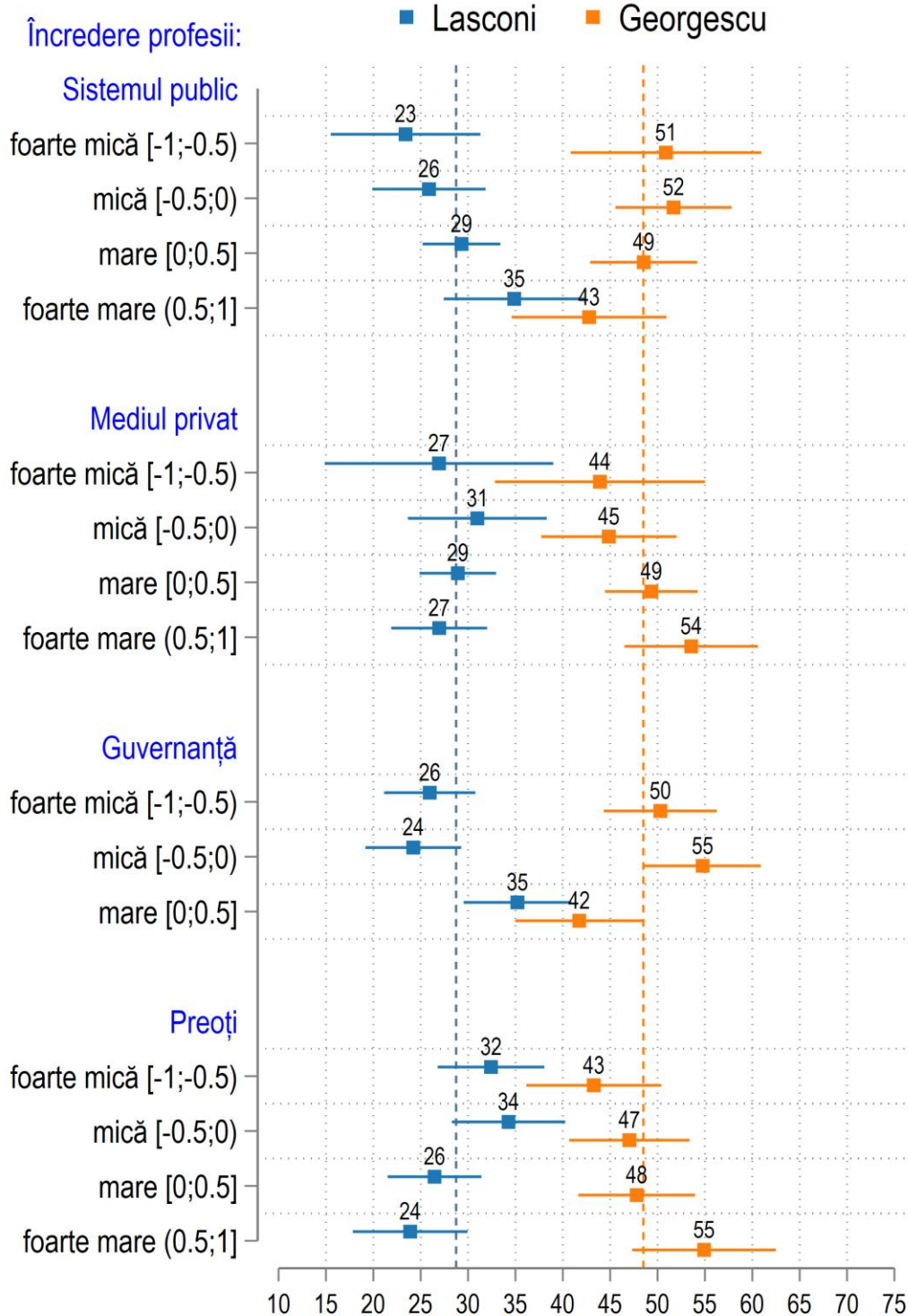


Cu cine ar fi votat la turul 2

■ Lasconi ■ Georgescu



Cu cine ar fi votat la turul 2



În urma analizei de mai sus am identificat **principali factori** (dintre cei pentru care am avut date) care, cu o mare probabilitate, **determină opțiunile de vot: nivelul de educație formală, evaluarea situației economice și politice din România în 2024 comparativ cu 2023, reprezentările cu privire la perioada comunistă și preferința pentru o societate de tip democratic.** În graficul de mai jos prezint valorile prezise ale intențiilor de vot la nivelul combinațiilor dintre clasele acestor variabile. Pe fiecare linie apar valorile prezise în cazul unei sub-populații, caracteristicile acesteia fiind precizate în primele patru coloane:

- educație (Edu): mai puțin decât liceul (<liceu), liceu/post-liceală (liceu), studii superioare (>liceu),
- evaluarea comparativă a situației economice și politice din România, 2024 vs. 2023 (RO): mai bine în 2023 ($23 > 24$), la fel ($23 = 24$), mai bine în 2024 ($23 < 24$),
- evaluarea comunismului (Com): bun (+) sau rău (-) și
- preferința pentru democrație (Demo): cel mult neutră (0), mare (+) și foarte mare (++)

Următoarea coloană indică ponderea opțiunilor de vot pentru Lasconi, respectiv Georgescu, iar ultima coloană indică diferența dintre cele două opțiuni. Observăm că intenția de vot pentru Georgescu este maximă în cazul combinațiilor: educație elementară sau liceu + situația din România s-a deteriorat + comunismul e bun + preferință neutră sau mare pentru democrație. Aproape în oglindă, intenția de vot pentru Lasconi este maximă în cazul combinațiilor: educație superioară + comunismul e rău + preferință foarte mare pentru democrație.

În concluzie, **principalii determinanți ai opțiunilor de vot** au fost: nivelul de educație formală, percepția cu privire la direcția în care s-a schimbat situația din România (economia, politica internă și politica externă), reprezentarea perioadei comuniste și preferința pentru o societate de tip democratic. **Cei cu studii superioare, care evaluează negativ perioada comunistă și pozitiv democrația ar fi votat dominant cu Lasconi. Cei care au absolvit cel mult liceul, care evaluează pozitiv perioada comunistă și nu preferă neapărat să trăiască într-o societate de tip democratic ar fi votat dominant cu Georgescu.**

Edu	RO	Com	Demo	Lasconi	Georgescu	Georgescu - Lasconi
< liceu	23>24	-	0	14	69	
< liceu	23>24	-	+	19	66	
< liceu	23>24	-	++	29	52	
< liceu	23>24	+	0	9	77	
< liceu	23>24	+	+	12	75	
< liceu	23>24	+	++	20	62	
< liceu	23=24	-	0	19	58	
< liceu	23=24	-	+	25	56	
< liceu	23=24	-	++	37	40	
< liceu	23=24	+	0	12	68	
< liceu	23=24	+	+	17	66	
< liceu	23=24	+	++	27	51	
< liceu	23<24	-	0	21	53	
< liceu	23<24	-	+	27	51	
< liceu	23<24	-	++	38	36	
< liceu	23<24	+	0	14	63	
< liceu	23<24	+	+	19	61	
< liceu	23<24	+	++	28	46	
liceu	23>24	-	0	13	63	
liceu	23>24	-	+	18	61	
liceu	23>24	-	++	27	46	
liceu	23>24	+	0	8	71	
liceu	23>24	+	+	12	70	
liceu	23>24	+	++	19	56	
liceu	23=24	-	0	18	52	
liceu	23=24	-	+	24	50	
liceu	23=24	-	++	34	35	
liceu	23=24	+	0	12	61	
liceu	23=24	+	+	16	60	
liceu	23=24	+	++	25	45	
liceu	23<24	-	0	19	46	
liceu	23<24	-	+	26	45	
liceu	23<24	-	++	35	30	
liceu	23<24	+	0	13	56	
liceu	23<24	+	+	18	55	
liceu	23<24	+	++	26	40	
> liceu	23>24	-	0	27	44	
> liceu	23>24	-	+	35	41	
> liceu	23>24	-	++	46	27	
> liceu	23>24	+	0	19	54	
> liceu	23>24	+	+	25	52	
> liceu	23>24	+	++	36	37	
> liceu	23=24	-	0	34	33	
> liceu	23=24	-	+	42	31	
> liceu	23=24	-	++	53	19	
> liceu	23=24	+	0	25	43	
> liceu	23=24	+	+	32	41	
> liceu	23=24	+	++	43	27	
> liceu	23<24	-	0	35	29	
> liceu	23<24	-	+	44	27	
> liceu	23<24	-	++	54	16	
> liceu	23<24	+	0	26	38	
> liceu	23<24	+	+	33	36	
> liceu	23<24	+	++	44	23	

Evident, „suspecții” incluși în aceste analize sunt doar o parte dintre factorii care pot explica opțiunile de vot. Am avut în vedere o serie de alți factori. Pentru a nu crește și mai mult durata medie de aplicare a chestionarului nu a fost posibilă includerea și mai multor întrebări. Oricum, au fost adresate mai multe întrebări decât ar fi fost recomandat pentru o anchetă telefonică. Cu și mai multe întrebări, calitatea datelor ar fi scăzut.

Toate valorile prezentate în graficele de mai sus trebuie privite cu anumite rezerve. Spun asta pentru că, având în vedere tipul de analize realizate, eșantionul a fost mai degrabă mic (964). Simplu spus, în cazul unora dintre combinațiile de variabile volumul eșantionului a fost mic. În astfel de situații este mai probabil ca eventualele cazuri puternic particulare pe alte dimensiuni să aibă un impact disproporționat asupra estimărilor produse în cazul combinației respective de variabile (caracteristici).

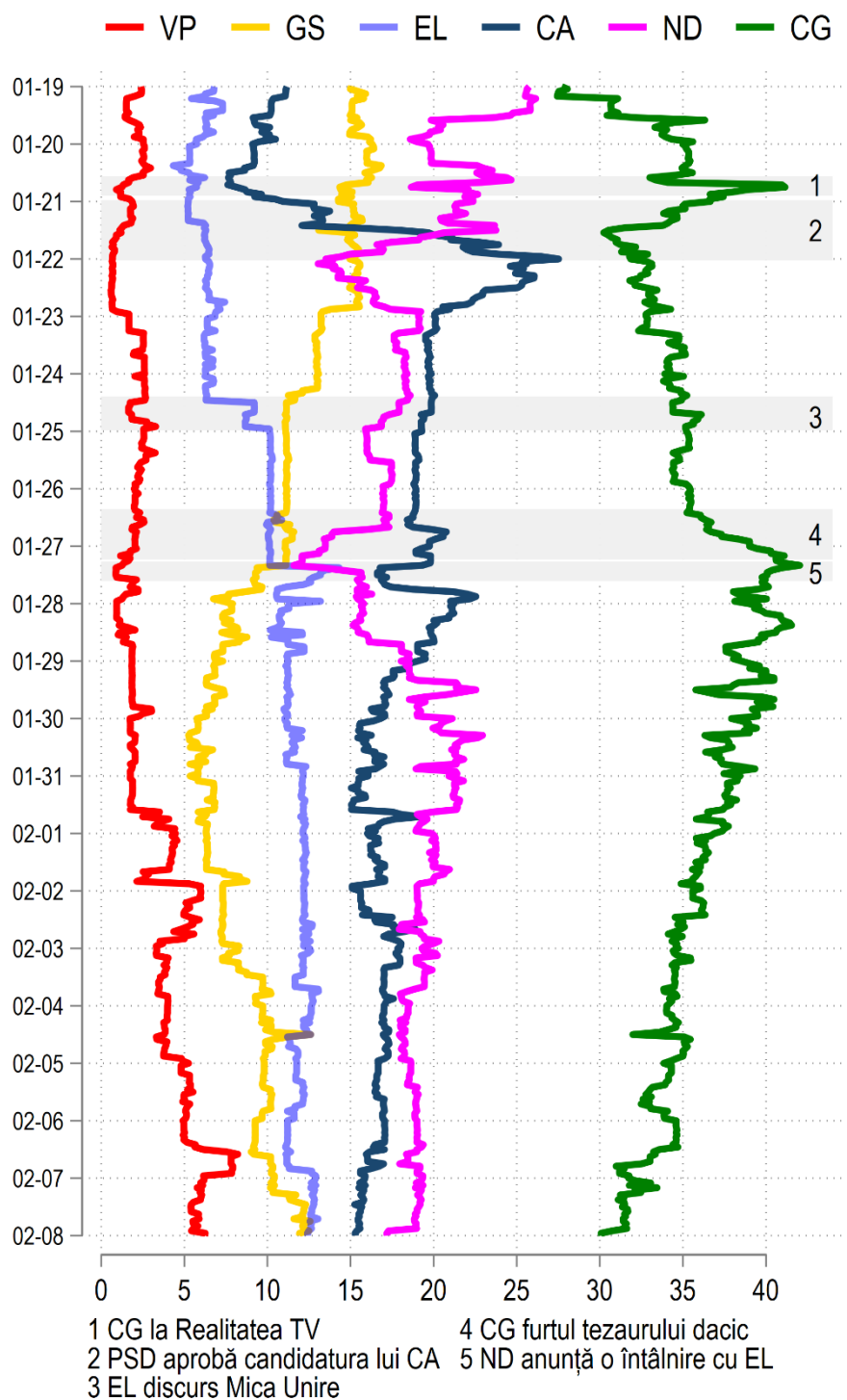
Prezidențiale 2025: cote pariuri (1)

Data: 2025-02-08

Cuvinte cheie: alegerile-prezidentiale-2025, cote-pariuri

Până apar mai multe sondaje, ne mulțumim cu cotele de la pariuri, transformate în procente, fără prea multe vorbe de astă dată!

Prezidențiale 2025: Cote pariuri



Prezidențiale 2025: cote pariuri (2)

Data: 2025-02-11

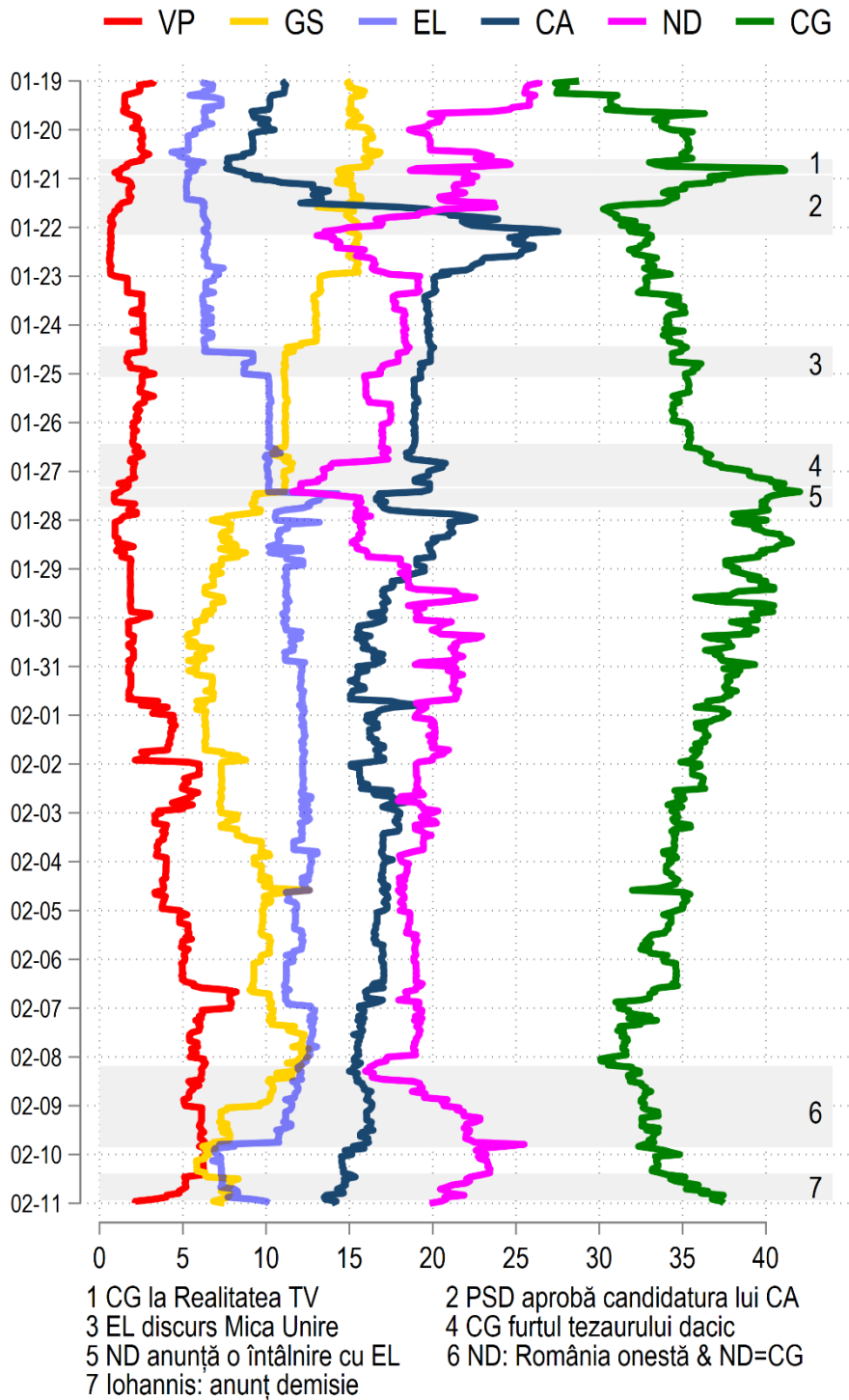
Cuvinte cheie: alegerile-prezidentiale-2025, cote-pariuri

Nici nu s-a uscat prea bine cerneala de la postarea anterioară și sunt nevoit să refac graficul construit pe baza cotelor de la pariuri pentru a surprinde (potențialul) impact al ultimelor evenimente: lansarea campaniei „România onestă” (Nicușor Dan) și anunțarea propriei demisii de către președintele Iohannis. Reamintesc că sondajele și casele de pariuri măsoară lucruri parțial diferite: sondajele adresează întrebarea „Dacă duminica viitoare ar fi alegeri prezidențiale, cu cine ați vota dintre ...?” iar casele de pariuri întrebarea „Care dintre următorii candidați credeți că va câștiga alegerile prezidențiale?”

E probabil ca anunțul cu privire la demisie să fi crescut șansele lui Călin Georgescu, respectiv să fi scăzut șansele lui Nicușor Dan și Victor Ponta. Spun „probabil” pentru că nu am controlat impactul posibil al altor factori și pentru că schimbările din ultimele ore depind parțial de schimbările anterioare, includ și o reacție a pieței la aceste schimbări, o încercare de a restabili echilibrul. Astfel, deoarece cota lui Nicușor Dan a crescut foarte mult în cele două zile anterioare anunțului președintelui Iohannis, o parte din scăderea de ieri poate fi produsă de tendința spre echilibru a pieței.

În ultimele date a apărut un nou candidat cu șanse mari. Deocamdată nu l-am inclus în analiză. Să vedem dacă efectul se menține măcar câteva zile.

Prezidențiale 2025: Cote pariuri



Clickbait sau lipsă de informare?

Data: 2025-02-18

Cuvinte cheie: alegerile-prezidentiale-2025, cote-pariuri

Tocmai am dat peste un articol publicat de Libertatea: [O companie internațională de predicții anunță că este posibil să nu avem un tur II al Alegerilor Prezidențiale. Cine conduce în sondaje?](#). Nu pare un articol de publicitate sau în scop de SEO pentru că are semnătura unui autor real de la redactorul Libertatea. Iată câteva afirmații din acest articol:

„... o companie internațională de predicții și unul dintre cei mai mari actori pe piața analizelor din mai multe domenii, a analizat situația politică din România și anunță că este posibil ca românii să voteze o singură dată și în această primăvară.”

„Sondajele demarate de PolyMarket îl clasează pe Călin Georgescu pe primul loc, cu o majoritate de 53%. În cazul în care rezultatele se vor transpune în realitate, Alegerile Prezidențiale din primăvara anului acesta vor fi câștigate dintr-un singur tur, românii fiind nevoiți să iasă la urne doar o dată.”

„Compania de sondaje a avut predicții apropiate de realitate și în cazul alegerilor anterioare din România, prezentând corect clasamentul alegerilor românilor atât la Alegerile Europarlamentare cât și la cele Locale din luna iunie 2024.”

Nu-mi dau seama dacă titlul și conținutul textului încearcă să atragă atenția cu orice preț sau dacă e vorba doar despre o lipsă de informare. Indiferent care dintre aceste două ipoteze este adevărată ... tot e aiurea. Iată de ce!

PolyMarket este o casă de pariuri, nu de sondaje, deci nu face sondaje. Să prezinți cotele de la pariuri ca rezultate ale unui sondaj ...!

PolyMarket nu face predicții. Desigur, datele respective pot fi și sunt folosite pentru a face predicții. Doar că pentru asta e nevoie de o modelare statistică a acestor date, alături de alte date.

PolyMarket (sau altă companie similară) nu face analize politice. Nici a situației din România, nici din alte țări. Cotele sunt stabilite automat, pe baza unui algoritm aplicat sumelor pariate.

Chiar dacă unul dintre candidații la alegerile prezidențiale din România obține mai mult de 50%+1 dintre voturi, probabilitatea ca acel candidat să câștige alegerile din primul tur e infimă. Argumentul este următorul:

- (1) Conform listelor electorale, în România sunt aproximativ 18,2 milioane persoane cu drept de vot.
- (2) În Constituția României scrie foarte clar (articolul 81, punctul 2): *Este declarat ales candidatul care a întrunit, în primul tur de scrutin, majoritatea de voturi ale alegătorilor înscriși în listele electorale.*
- (3) La alegerile prezidențiale, de altfel cele cu prezența cea mai mare, s-au prezentat până acum la turul 1 undeva în jur de 10 milioane de alegători.

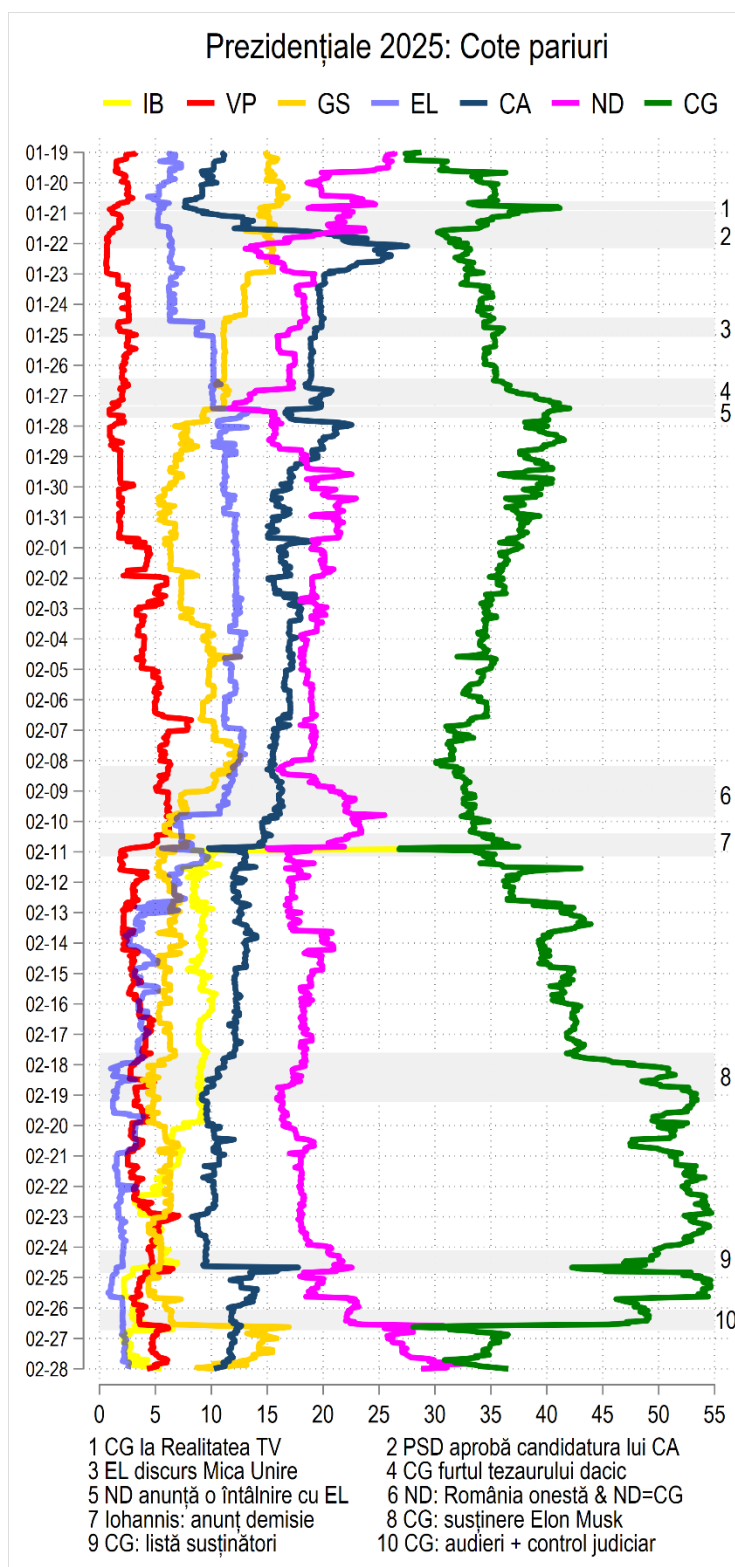
Din aceste trei premise rezultă că un candidat poate deveni președinte din primul tur doar dacă obține aproximativ 9,1 milioane voturi, adică peste 90% din voturile care de obicei sunt exprimate la turul 1. Crede cineva că un astfel de rezultat are o probabilitate cât de cât rezonabilă? Să omiți un fapt atât de important, nu prea seamănă cu jurnalismul pe care ni-l dorim, nu?!

Prezidențiale 2025: cote pariuri (3)

Data: 2025-02-28

Cuvinte cheie: alegerile-prezidentiale-2025, cote-pariuri

Un update cu cotele de la pariuri în urma ultimelor evenimente. Pe lista candidaților cu un scor ridicat apare și Ilie Bolojan.



O estimare într-un context electoral marcat de incertitudini

Data: 2025-03-01

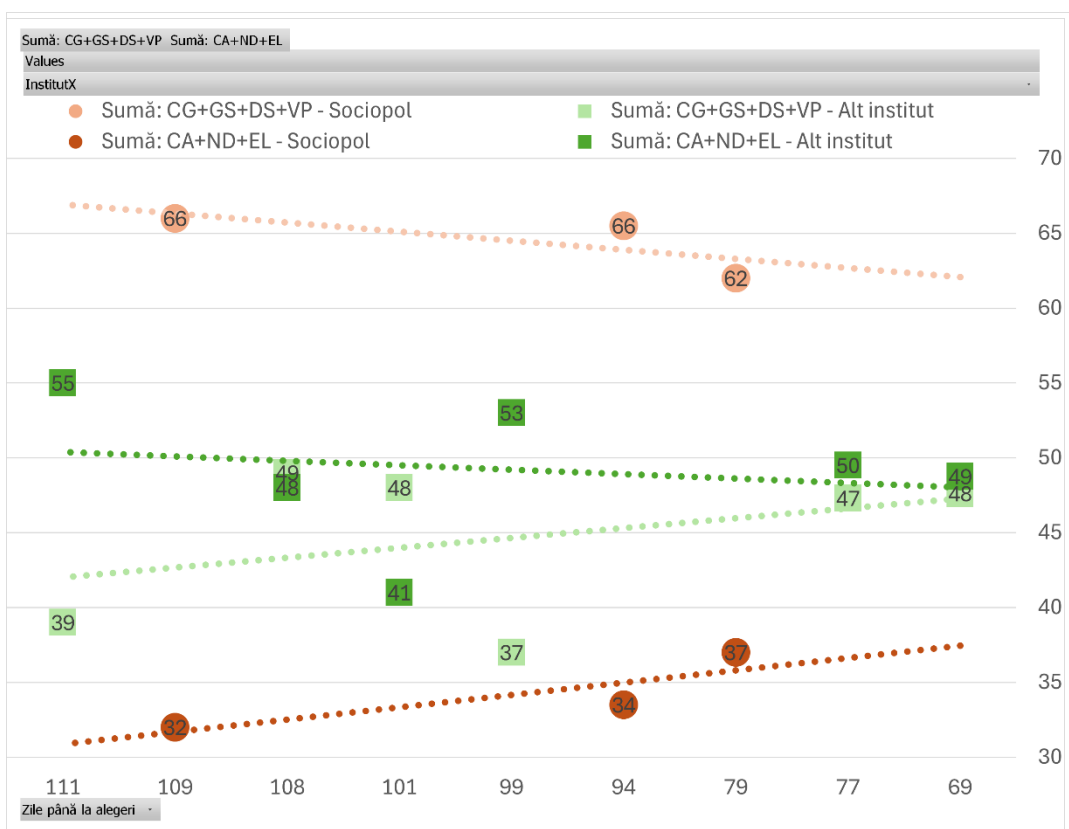
Cuvinte cheie: alegerile-prezidentiale-2025, dinamica-intentiilor-de-vot

Până în prezent, anul acesta au fost publicate foarte puține sondaje de opinie pe tema alegerilor prezidențiale din mai 2025. În acest moment, lista candidaților este mai degrabă orientativă, fiind posibil să asistăm în următoarele două săptămâni la intrări și ieșiri importante. Institutele de sondare au testat diferite combinații de candidați, nu neapărat aceleași de fiecare dată. Ca de obicei, unele sondaje au fost realizate telefonic (CATI), altele online. Toate aceste lucruri au condus la estimări destul de diferite ale intențiilor de vot de la un sondaj la altul. Astfel, în cazul unor candidați apar diferențe mari, uneori chiar de 10-15 puncte procentuale. În astfel de condiții, o analiză longitudinală a intențiilor de vot la nivel de candidați este sortită din start eșecului. Se poate face însă o analiză la nivel agregat, prin însumarea intențiilor de vot pentru candidații relativ apropiați din punctul de vedere al alegătorilor. Desigur, orice agregare este discutabilă. Ea poate fi însă utilă pentru a estima, într-un context marcat de incertitudine, fie și grosier, tendința generală a preferințelor de vot.

Agregarea folosită aici este următoarea: **Georgescu + Simion + Șoșoacă + Ponta vs. Antonescu + Dan + Lasconi**. O astfel de agregare permite, cu anumite limite, compararea intențiilor de vot pentru fiecare dintre cele două poluri care par să definească spațiul alegerilor prezidențiale. Suplimentar, pentru că

diferențele dintre estimările produse de diferite institute de sondare sunt mari, am prezentat rezultatele agregării în funcție de institutul care a realizat ancheta: **Sociopol** vs. **Alt institut** (apar diferențe și între aceste institute, dar nu atât de mari). Valorile prezentate în grafic sunt medii ale estimărilor produse de toate variantele testate de sondajele publicate până în acest moment. Pentru că numărul de sondaje este foarte mic și scenariile testate diferă adesea, estimările sunt mai degrabă orientative.

Funcție de institut, cele două poluri diferă ca pondere și dinamică în timp. În sondajele Sociopol polul **Georgescu + Simion + Șoșoacă + Ponta** are în prezent o pondere de peste 60%, dinamica fiind ușor descendentă, iar polul **Antonescu + Dan + Lasconi** are o pondere de aproximativ 35% și o dinamică ușor ascendentă. În restul sondajelor, cele două poluri au în prezent aproximativ aceeași pondere (~48%), în acest caz dinamica polului **Georgescu + Simion + Șoșoacă + Ponta** fiind ușor ascendentă, iar a polului **Antonescu + Dan + Lasconi** ușor descendentă.



O estimare tot la nivel agregat

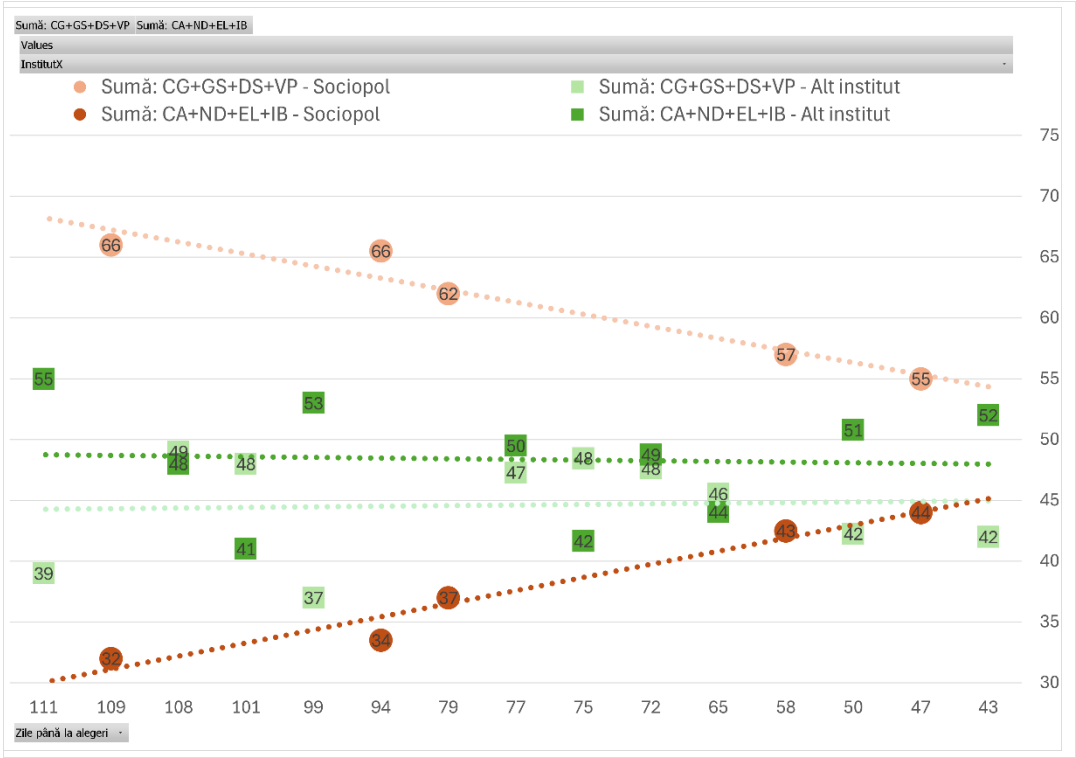
Data: 2025-03-23

Cuvinte cheie: alegerile-prezidentiale-2025, dinamica-intentiilor-de-vot

În postarea [anterioară](#) am prezentat un grafic cu o estimare longitudinală, la nivel agregat, a intențiilor de vot pentru alegerile prezidențiale din mai. Între timp au mai apărut câteva sondaje, prin urmare refac aici graficul. În continuare, estimările intențiilor de vot diferă destul de mult de la un sondaj la altul. Pare să existe totuși o tendință de convergență. Reamintesc agregarea folosită: **Georgescu + Simion + Șoșoacă + Ponta vs. Antonescu + Dan + Lasconi + Bolojan**. Pentru că diferențele dintre estimările produse de diferite institute de sondare sunt mari, am prezentat rezultatele agregării în funcție de institutul care a realizat ancheta: **Sociopol vs. Alt institut** (apar diferențe și între aceste institute, dar nu atât de mari). Valorile prezentate în grafic sunt medii ale estimărilor produse de toate variantele testate de sondajele publicate până în acest moment.

Funcție de institut, cele două poluri diferă ca pondere și dinamică în timp. Astfel, în sondajele Sociopol, polul **Georgescu + Simion + Șoșoacă + Ponta** are în prezent o pondere de aproximativ 55%, dinamica fiind descendentă (scădere de ~10 puncte procentuale), iar polul **Antonescu + Dan + Lasconi + Bolojan** are o pondere de aproximativ 45% dinamica fiind ascendentă (creștere de ~10 pp). În restul sondajelor, cele două poluri au în prezent aproximativ aceeași pondere (48%, respectiv 45%), în ambele cazuri schimbarea fiind absentă. Parțial, absența unei schimbări este rezultatul faptului că încep să apară diferențe relativ mai mari și

între sondajele publicate de restul institutelor, nu doar între Sociopol și alte institute.



„House effects”: alegerile prezidențiale 2025, T1 (1)

Data: 2025-03-26

Cuvinte cheie: alegerile-prezidentiale-2025, house-effects

Până în prezent au fost publicate 10 sondaje (6 Sociopol, 4 alte institute) în care apar principalii candidați care s-au și înscris în cursa pentru alegerile prezidențiale din 4 mai 2025. Chiar dacă numărul de sondaje e foarte mic, pare că deja pot fi observate efectele de institut (house-effect). În condițiile prezente, o estimare a efectelor e doar orientativă, pentru următoarele motive:

- numărul total de sondaje e foarte mic,
- numărul de sondaje realizate în aproximativ aceeași perioadă e foarte mic,
- unele institute au publicat un singur sondaj,
- efectul de metodă (CATI vs. Online) se suprapune parțial cu efectul de institut și,
- **nu știm rezultatele finale, deci e posibil ca o deviere de la medie să reflecte de fapt o estimare mai corectă.**

Graficul de mai jos prezintă estimarea intențiilor de vot, comparativ între institute. Valorile respective sunt valori prezise pe baza unui model de regresie logistică multi-nominală. Modelul controlează influența perioadei în care au fost culese datele și tehnica de ancheta (CATI vs Online). Simbolurile reprezintă estimările asociate candidaților, iar liniile intervalul de încredere.

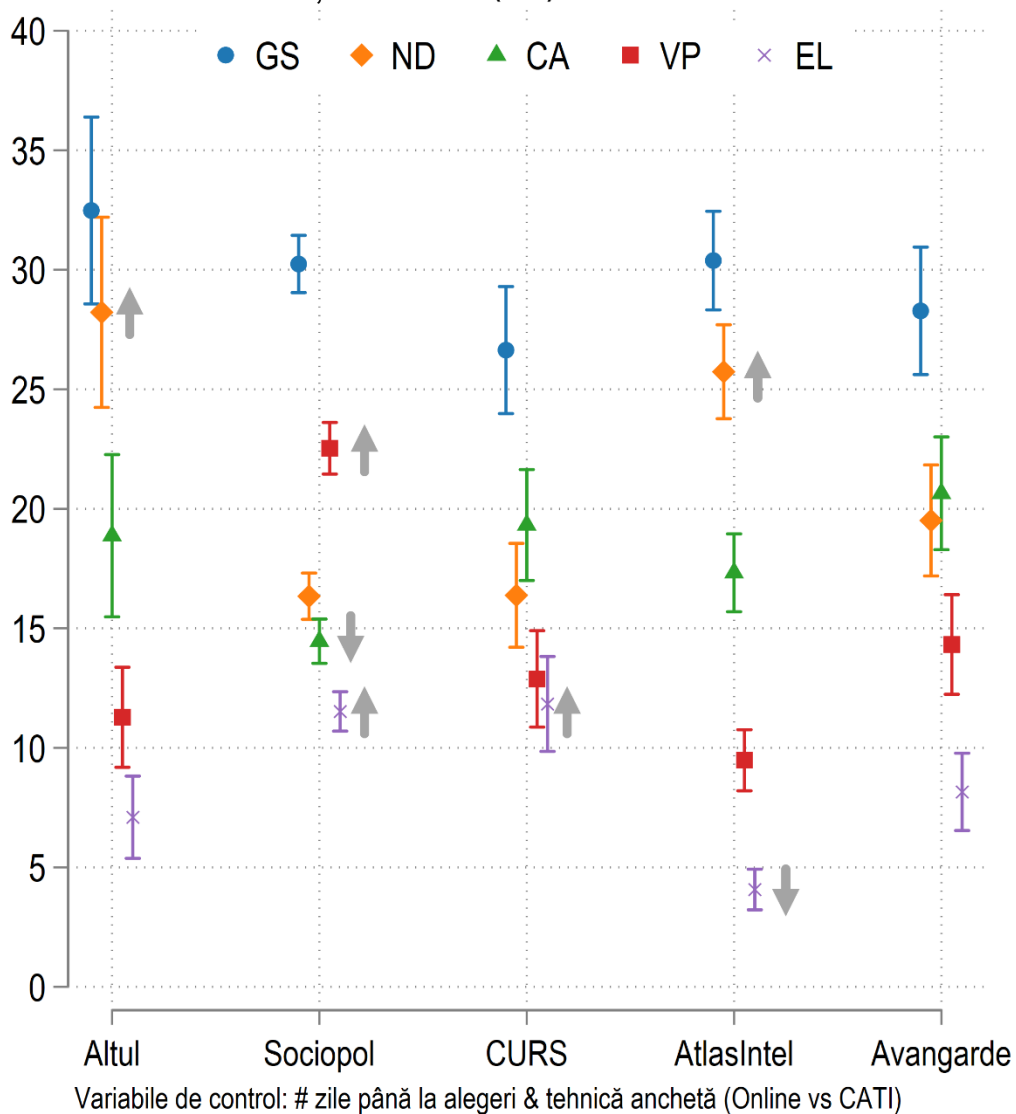
Săgețile indică sensul diferenței observate: dacă acel institut a publicat, comparativ cu tendința generală, o estimare mai mare / mică în cazul unui anumit candidat. **Valorile din grafic nu reprezintă estimări ale intențiilor de vot de la un moment dat, nici estimări ale procentelor care vor fi obținute de candidați la alegeri. Valorile reprezintă predicțiile realizate pe baza unui model în care am controlat influența a două variabile (numărul de zile până la alegeri și tehnica de anchetă) cu scopul de a vedea dacă mai rămân diferențe între estimările produse de către institutele de sondare.**

La ce e utilă o astfel de analiză? E normal să apară diferențe între estimările produse de diferite institute. Firesc, ne putem întreba cât de mari trebuie să fie aceste diferențe pentru a le considera puțin probabile din punct de vedere statistic. Diferențele dintre estimările diferitelor institute, așa cum sunt ele observate „cu ochiul liber,” pot avea ca sursă o serie de factori absolut legitimi precum perioada de culegere a datelor și tehnica de anchetă. Prin modelare, am eliminat impactul posibil al acestor factori. Dacă observăm că institutele produc estimări diferite chiar și atunci când controlăm statistic factorii menționați, putem spune că diferențele respective au o altă sursă decât perioada de colectare a datelor și tehnica de anchetă (CATI vs. Online).

Câteva concluzii cu privire la estimările intențiilor de vot, pe scurt:

- **George Simion:** nu apar diferențe notabile între institute (intervalele de încredere asociate estimărilor se suprapun măcar parțial);
- **Nicușor Dan:** estimări mai mari în cazul institutelor AtlasIntel și Altul (Noi, cetățenii); ambele anchete au fost online;
- **Crin Antonescu:** estimare mai mică în cazul Sociopol;
- **Victor Ponta:** estimare mai mare în cazul Sociopol;
- **Elena Lasconi:** estimare mai mare în cazul Sociopol și CURS, respectiv mai mică în cazul AtlasIntel.

Prezidențiale 2025 (T1): House effects



Dinamica intențiilor de vot: turul 1 al alegerilor prezidențiale din mai 2025 (1)

Data: 2025-03-27

Cuvinte cheie: alegerile-prezidentiale-2025, dinamica-intențiilor-de-vot

Cine se uită la estimările intențiilor de vot conform sondajelor publicate are șanse mari să devină tot mai sceptic cu privire la corectitudinea acestora. Dacă avem în vedere ultimele patru sondaje publicate, realizate aproximativ în aceeași perioadă și cu aceeași metodologie (sondajul AtlasIntel a fost online - RDR, restul CATI), observăm că ierarhiile sunt foarte diferite, cu excepția candidatului care ocupă prima poziție (vezi și [postarea de ieri](#)). În aceste condiții, nu e de mirare că tot mai mulți alegători percep sondajele mai degrabă ca un mijloc de manipulare și nu unul de informare.

Diferențele dintre sondaje pot fi explicate prin diferiți factori de natură tehnică. Unii factori au legătură cu metodologia de colectare a datelor: tehnica utilizată (CATI sau Online), chestionarul (formularea și ordinea întrebărilor) și eșantionarea (numere de telefon selectate aleator, panel online etc.). Un alt factor este numărul redus de cazuri: la un eșantion de 1000 persoane, dacă doar jumătate s-au decis și au declarat cu cine votează, intențiile de vot sunt calculate în realitate pe 500 cazuri, nu pe 1000, deci marja de eroare e mai mare decât cea comunicată (+/-3 puncte procentuale, cel mai adesea). Prin urmare, o estimare a intenției de vot de 30% poate fi în realitate undeva în

intervalul 26-34%, nu 27-33% (intervalul teoretic e în realitate mai mic, dar nu mai stau acum să-l calculez exact). Și asta doar dacă (1) eșantionul este unul simplu aleator, ceea ce nu e cazul aproape niciodată (nici la start și, mai ales, nici la final) și (2) nu apar erori de alt tip, care nu se anulează între ele, ceea ce evident este din nou fals cel mai adesea. Mai multe explicații cu privire la astfel de probleme pot fi citite [aici](#).

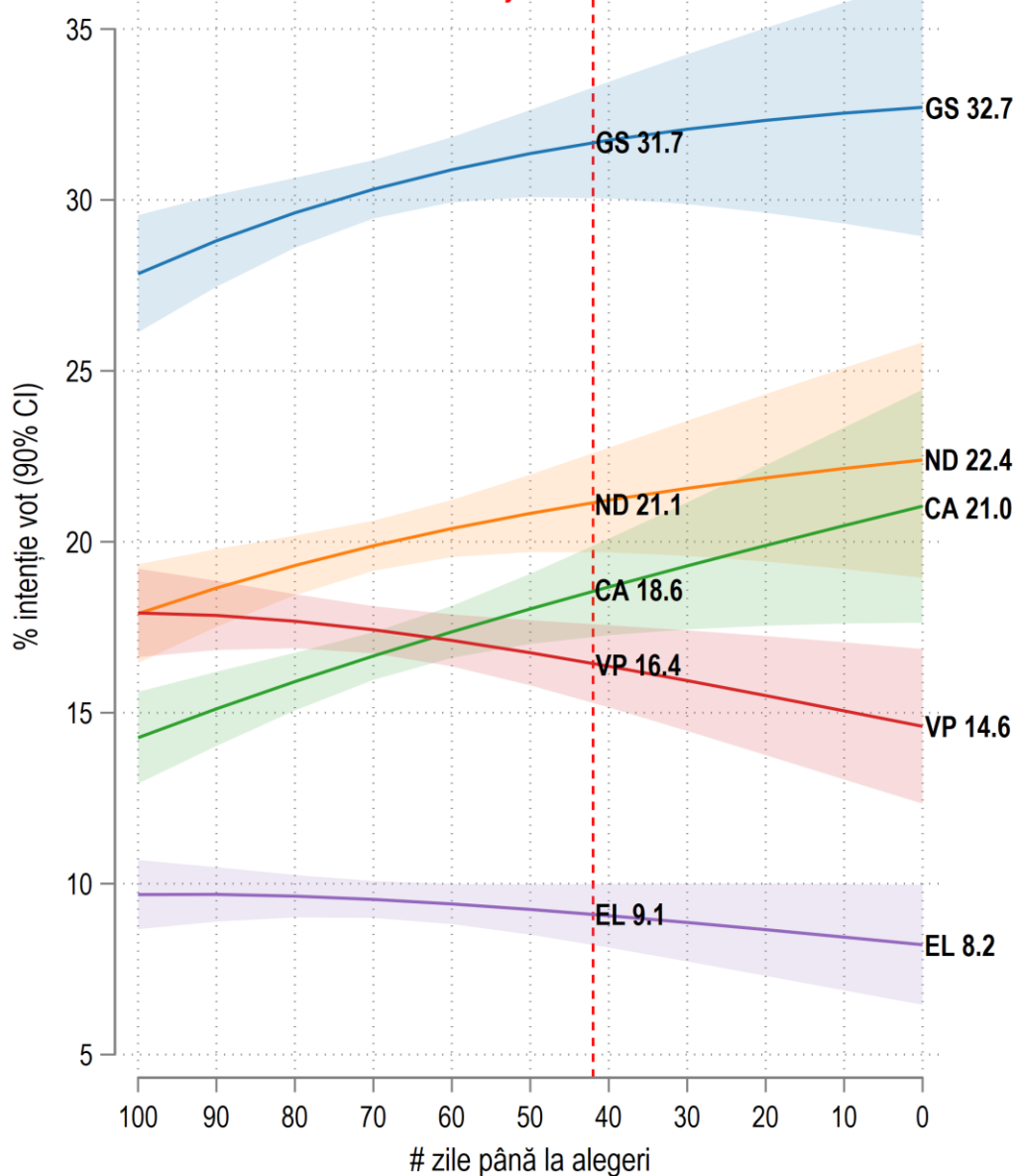
Înainte de a privi graficul, aveți în vedere următoarele:

- estimările din grafic sunt produse pe baza estimărilor din cele 10 sondaje publicate până acum (6 Sociopol, 4 alte institute); doar în aceste sondaje apar principalii candidați care s-au și înscris în cursa pentru alegerile prezidențiale din 4 mai;
- în prima parte a perioadei analizate, majoritatea sondajelor publicate au fost realizate de un singur institut (Sociopol); o astfel de situație poate influența mult estimările și tendințele de vot observate;
- estimările din grafic sunt predicții realizate folosind un model de regresie logistică multi-nominală; variabila independentă de interes: numărul de zile până la alegeri; variabile de control: Institutul și tehnica de anchetă (CATI vs. Online).
- estimările se referă la intențiile de vot de pe teritoriul României (doar AtlasIntel colectează date și din diaspora, din câte știu);
- aceste estimări nu iau în calcul calitatea estimărilor produse anterior de institute și nici eventualul bias al acestor institute;
- estimările nu țin cont de posibilele efecte induse de (in)dezirabilitatea socială și/sau ratele diferite de participare la sondaje asociate electoratului diferiților candidați (astfel de factori pot influența toate estimările în același sens);

- diferența până la 100% indică intențiile de vot pentru restul candidaților;
- graficul prezintă două estimări pentru fiecare candidat: (1) estimarea din proximitatea liniei verticale întrerupte de culoare roșie ne arată **situația probabilă de acum** (mai exact situația din ultima zi de teren a celui mai recent sondaj – 23 martie 2025); (2) estimarea din dreapta graficului indică **rezultatul probabil de la alegeri** pe baza sondajelor existente la acest moment, asumând că tendința observată va continua și că perioada de campanie nu va fi afectată de evenimente majore; evident, în cazul estimărilor pentru ziua votului incertitudinea asociată este mult mai mare (banda colorată asociată estimărilor devine tot mai lată pe măsură ce ne îndepărtăm de zilele pentru care există date de sondaj).

Prezidențiale 2025 (T1): Dinamica intențiilor de vot (valori prezise pe baza sondajelor publicate)

Ultima zi cu date de sondaj: **23mar2025**



Variabile de control: Institut & Tehnică anchetă (Online vs CATI)

Cât de „disciplinat” e electoratul partidelor?

Data: 2025-04-03

Cuvinte cheie: alegerile-prezidentiale-2025

Într-o reformulare ceva mai explicită, în această analiză îmi propun să identific (in)congruența dintre preferințele alegătorilor pentru partide și candidații susținuți de acestea în competiția electorală din mai 2025. Simplu spus: **Care este ponderea votanților unui partid care intenționează să (nu) voteze cu candidatul susținut de acel partid?** La acest moment, o astfel de analiză poate fi realizată doar pe baza datelor din sondaje, adică strict pe baza unor declarații ale posibililor votanți, de unde și o serie de limite.

Pentru a face o astfel de analiză trebuie să cunoaștem distribuția încrucișată a răspunsurilor alegătorilor la două întrebări de sondaj: (1) **Care este partidul votat la ultimele alegeri (Parlamentare 2024)?** și (2) **Care este intenția de vot pentru alegerile din mai?** Problema e că, majoritatea rapoartelor prezentate public de institutele de sondare prezintă doar distribuția răspunsurilor la fiecare întrebare în parte, nu și distribuția încrucișată. Suplimentar, pentru că un eșantion obișnuit are puțin peste 1000 respondenți, iar un tabel care încrucișează intenția de vot pentru președinte cu votul anterior are minimum 25 combinații, rezultă că o medie de cel mult 40 de respondenți pe combinație. Pentru că nu toți respondenții răspund la întrebările de vot, în realitate, numărul e mai mic. Estimările care au la bază un număr atât de mic de cazuri pot fi foarte diferite de valorile reale. Putem crește numărul de cazuri

doar dacă folosim mai multe sondaje. Chiar și așa, erorile pot fi mari în unele cazuri.

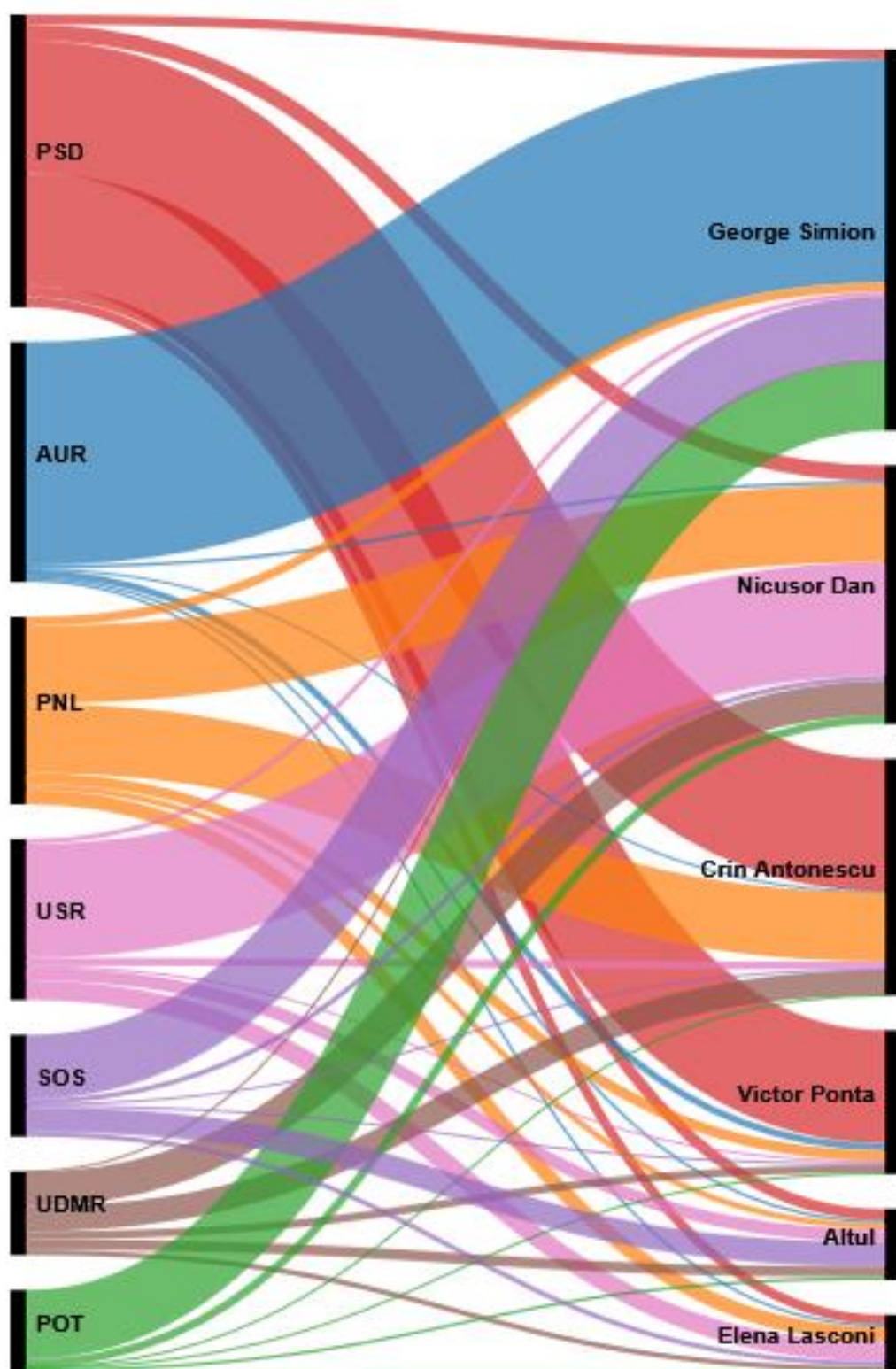
Am identificat două sondaje, ambele online, care au măsurat variabilele de interes și la datele cărora am avut acces. [Raportul AtlasIntel](#) din martie 2025 prezintă un tabel cu distribuția încrucișată a celor două variabile de interes. În cazul celei de a doua surse, [ultimul sondaj MKOR](#) (sfârșit de aprilie), am primit baza de date și am produs eu tabelul. Cum cele două surse prezentau date relativ similare, le-am combinat (am calculat mediile).

Pornind de la tabelul produs anterior și de la numărul de voturi primite de partide la alegerile parlamentare din 2024 am realizat graficul de mai jos. În stânga avem partidele parlamentare, iar în dreapta candidații la alegerile prezidențiale din acest an. Grosimea unei benzi colorate indică numărul de voturi primite de un partid, respectiv numărul de voturi ale acelui partid care merg, pe baza declarațiilor din sondaje, spre un anumit candidat. Pentru că estimările prezentate sunt calculate folosind și date de sondaj, ele trebuie privite mai degrabă ca tendințe, nu ca valori absolute. Observăm următoarele:

- Majoritatea voturilor primite de PSD se împart aproximativ egal între Crin Antonescu și Victor Ponta (45%, respectiv 39%).
- Majoritatea voturilor primite de PNL se împart aproximativ egal între Nicușor Dan și Crin Antonescu (41%, respectiv 37%).
- Cea mai mare parte a voturilor primite de UDMR se împart aproximativ egal între Nicușor Dan și Crin Antonescu (41%, respectiv 30%).
- Majoritatea voturilor primite de USR merg la Nicușor Dan (71%) și doar în jur de 12% la Elena Lasconi (probabil procentul real e ceva mai mare - numărul de cazuri e mic, sondajele sunt online).
- Aproape toate voturile primite de AUR merg la George Simion (93%).

- Majoritatea voturilor obținute de POT merg la George Simion (82%).
- Peste jumătate din voturile primite de S.O.S. merg la George Simion (63%).

Pe baza acestor estimări putem spune că: **majoritatea alegătorilor AUR și POT sunt disciplinați, adică votează candidatul AUR; cel mult jumătate din electoratele PSD, PNL și UDMR votează candidatul susținut de coaliția formată din aceste partide; electoratul USR este cel mai puțin disciplinat, votând majoritar un alt candidat decât cel susținut oficial de partid. În concluzie, pe baza tuturor datelor disponibile, putem aproxima că doar jumătate (52%) din electoratul partidelor din România este disciplinat, în sensul că votează cu candidatul susținut de propriul partid.**



Relația dintre încredere și vot

Data: 2025-04-05

Cuvinte cheie: alegerile-prezidentiale-2025, încredere, intentii-de-vot

În cadrul acestei analize, prin încredere mă refer la încrederea în diferite instituții și în oameni, iar prin vot la intenția de a vota cu un anumit candidat la alegerile prezidențiale din 4 mai 2025. Datele analizate aici provin din [ultimul sondaj MKOR](#) (mulțumesc pentru baza de date). Analizele, graficele și comentariile îmi aparțin.

Tabelul de mai jos prezintă comparativ încrederea medie în instituții a votanților principalilor candidați (AL = alt candidat), plus total votanți. Scala utilizată a fost 1-7, unde 1 înseamnă încrederea cea mai mică, iar 7 încrederea cea mai mare (4 = mijlocul scalei). Mărimea valorilor medii variază în funcție de candidat, dar mai ales în funcție de instituția evaluată. O singură valoare medie este mai mică de 2 și foarte puține sunt mai mari de 5. La un pol avem instituții care beneficiază de un **nivel foarte redus de încredere** (Guvern, Parlament, Partide), la celălalt pol instituții în care votanții au un **nivel intermediar de încredere** (Școală), restul instituțiilor beneficiind de un nivel redus de încredere.

Încredere (medie 1-7)	GS	ND	CA	VP	EL	AL	Total
Guvern	1.9	3.4	4.2	3.1	2.7	2.3	2.9
Parlament	2.7	3.0	3.7	2.9	2.8	2.2	3.0
Partide politice	2.1	2.9	3.4	2.6	2.4	2.1	2.6
Administrație publică locală	4.0	4.0	4.4	4.2	3.9	2.7	4.0
Justiție	2.7	4.2	4.3	3.6	4.4	2.7	3.5
Poliție	4.0	4.6	4.9	4.4	4.8	3.0	4.3
Sistem sanitar	4.2	4.6	4.7	4.0	4.2	3.2	4.3
Școală	5.5	4.9	5.1	5.1	5.3	3.7	5.1
Companii private	3.6	4.6	4.9	4.2	4.6	3.4	4.2
Presă	3.1	4.1	3.9	3.6	3.6	3.3	3.6
Bănci	3.7	4.1	4.1	4.4	4.0	3.4	4.0
BNR	4.0	4.9	5.5	4.7	4.1	3.7	4.6
Biserică	5.2	3.5	5.1	4.8	4.7	4.5	4.6
Alți oameni	3.2	4.0	4.5	4.0	4.6	3.7	3.9

Graficele de mai jos prezintă relația dintre preferința de vot și încredere așa cum este descrisă de rezultatele unui model de regresie logistică multinominală (variabile de control: sex, vârstă, educație și mediu de rezidență). Un astfel de grafic se citește astfel: celelalte condiții considerate fiind aceleași (sex, vârstă etc.), probabilitatea de vot pentru candidatul X crește/scade/rămâne constantă pe măsură ce încrederea în instituția Y crește (cele două fenomen corelează; strict pe baza acestor date și analize nu poate fi stabilită cu certitudine existența unei relații de tip cauzal între încredere și preferința de vot). De exemplu, dacă un alegător „mediu” are o încredere minimă în Guvern (1, pe o scală de la 1 la 7), probabilitatea acestuia de a vota cu George Simion este 45% (respectiv 8% să voteze cu Crin Antonescu); dimpotrivă, dacă un alegător „mediu” are o încredere maximă (7), probabilitatea de a vota cu Crin Antonescu este 45% (respectiv 5% să voteze cu George Simion).

În general, **probabilitatea de vot pentru Elena Lasconi** nu variază în funcție de încrederea în instituții. Observăm totuși o relație pozitivă în cazul încrederii în Justiție, Companiile private și alți oameni, prin urmare, probabilitate de vot pentru Elena Lasconi crește pe măsură ce crește încrederea în aceste instanțe.

În general, **probabilitatea de vot pentru Victor Ponta** variază puțin în funcție de încrederea în instituții. Observăm o relație pozitivă în cazul încrederii în Administrația publică locală, Bănci

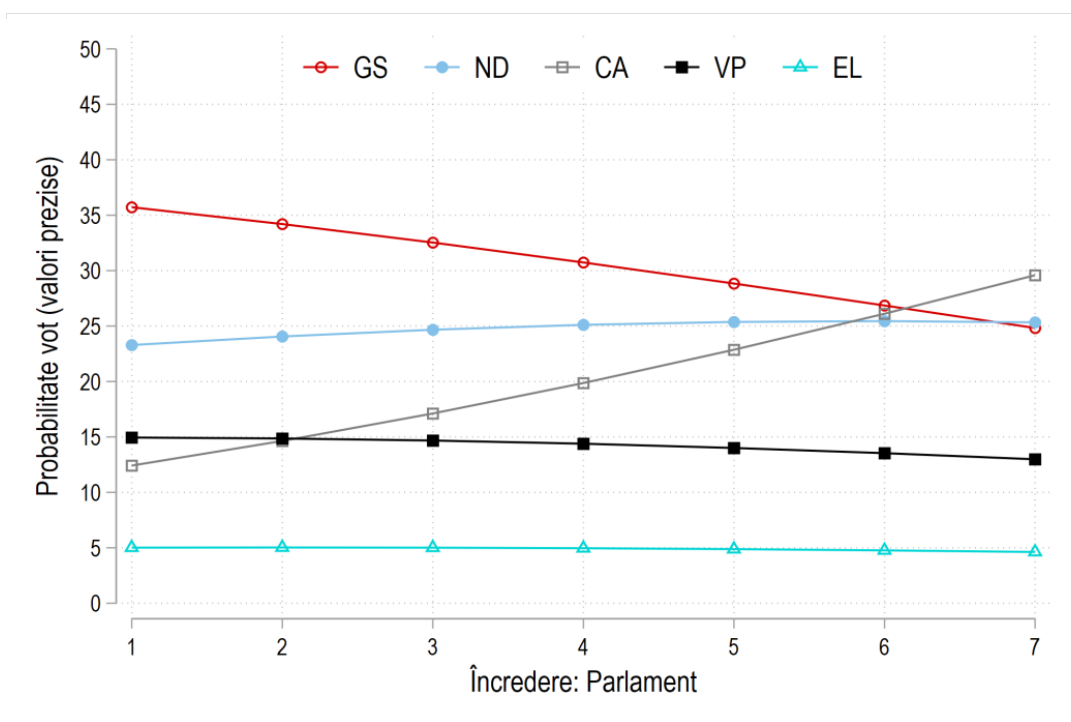
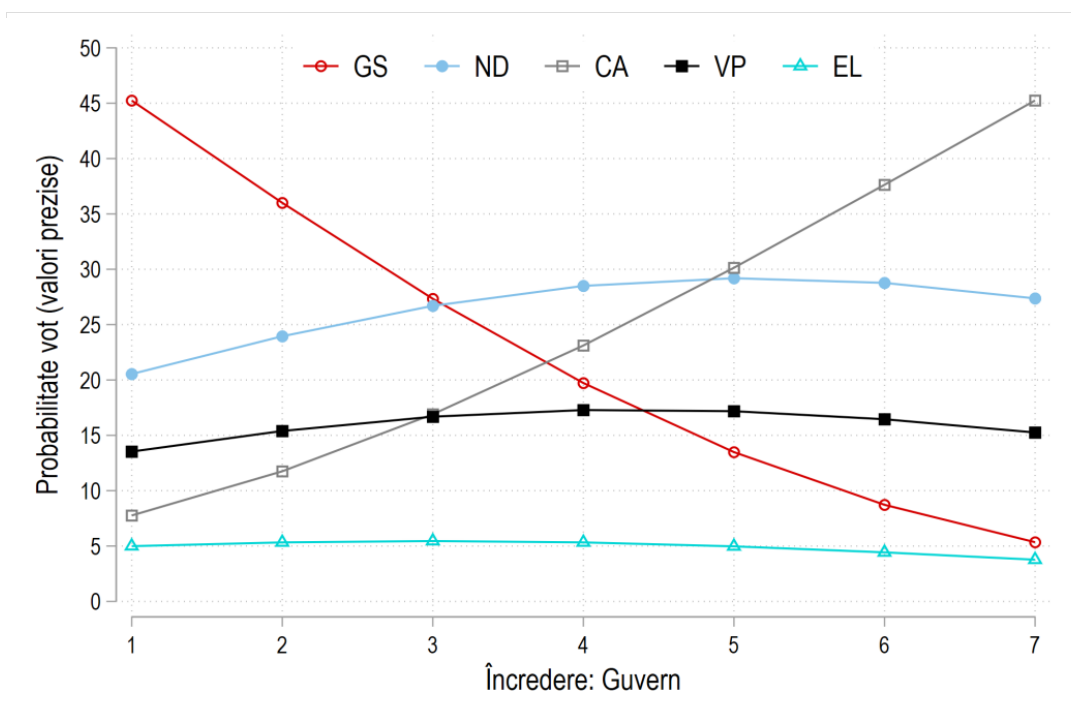
și Biserică. O relație negativă apare în cazul încrederii în Sistemul sanitar.

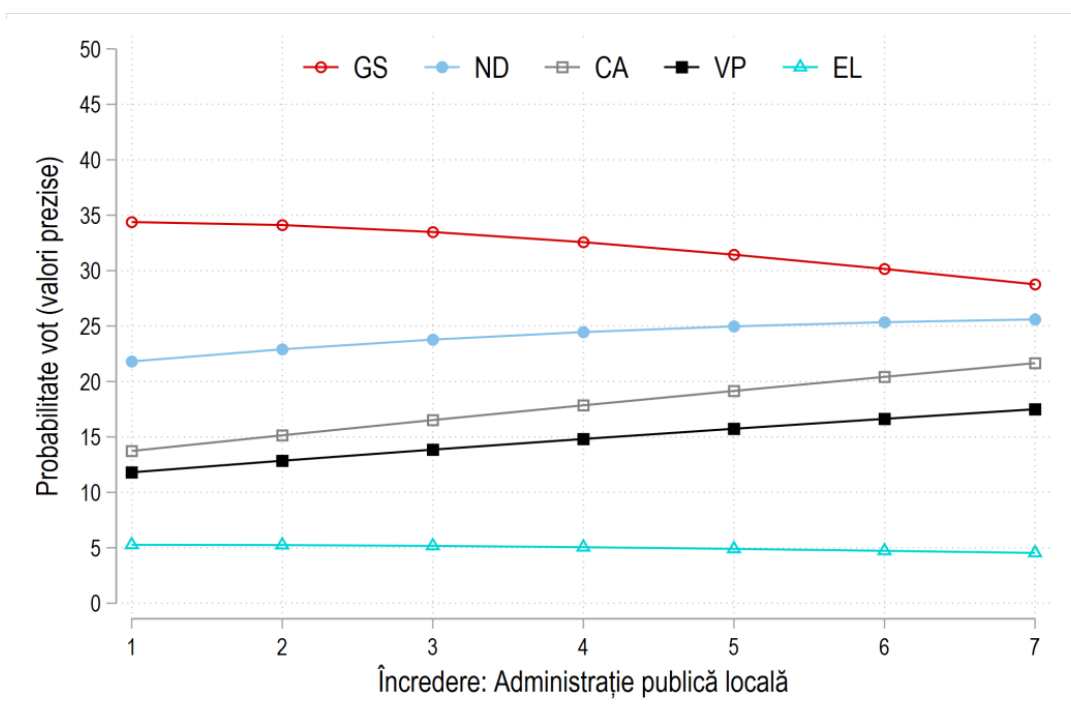
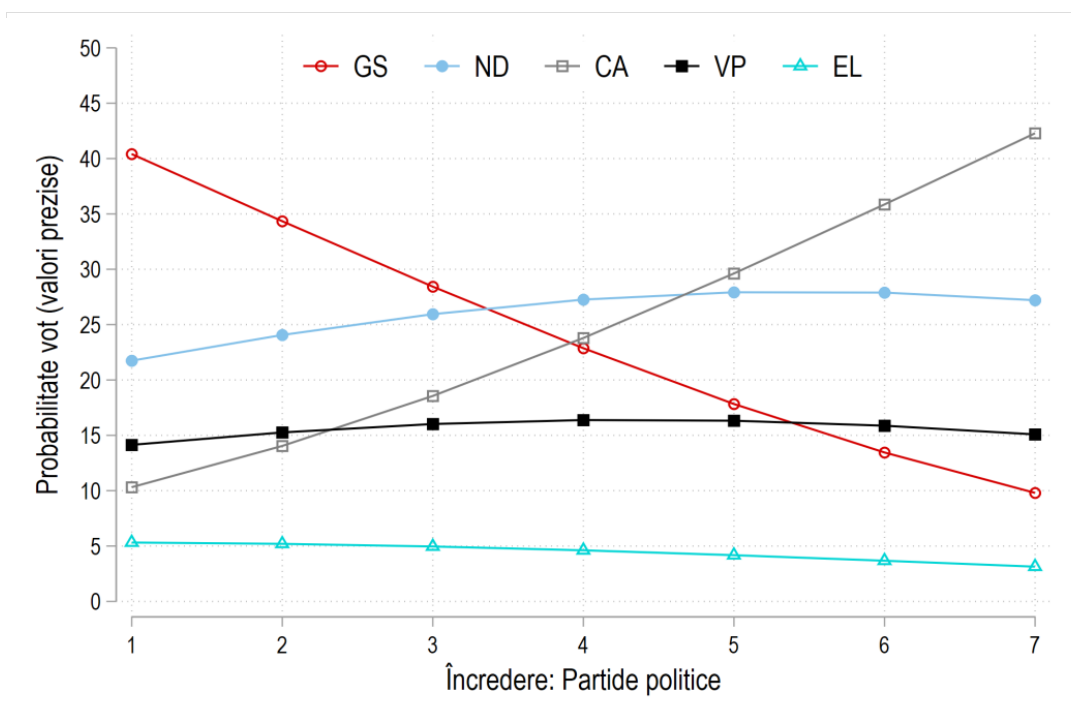
Probabilitatea de vot pentru Crin Antonescu variază destul de mult în funcție de încrederea în instituții. În doar câteva instanțe (Școală, Presă și Bănci) observăm lipsa unei relații. În restul situațiilor (mai ales Guvern, Parlament, Partide, Justiție, Companii private și alți oameni), observăm o relație pozitivă, adică, cu cât încrederea este mare, cu atât probabilitatea de vot pentru Crin Antonescu este mai mare.

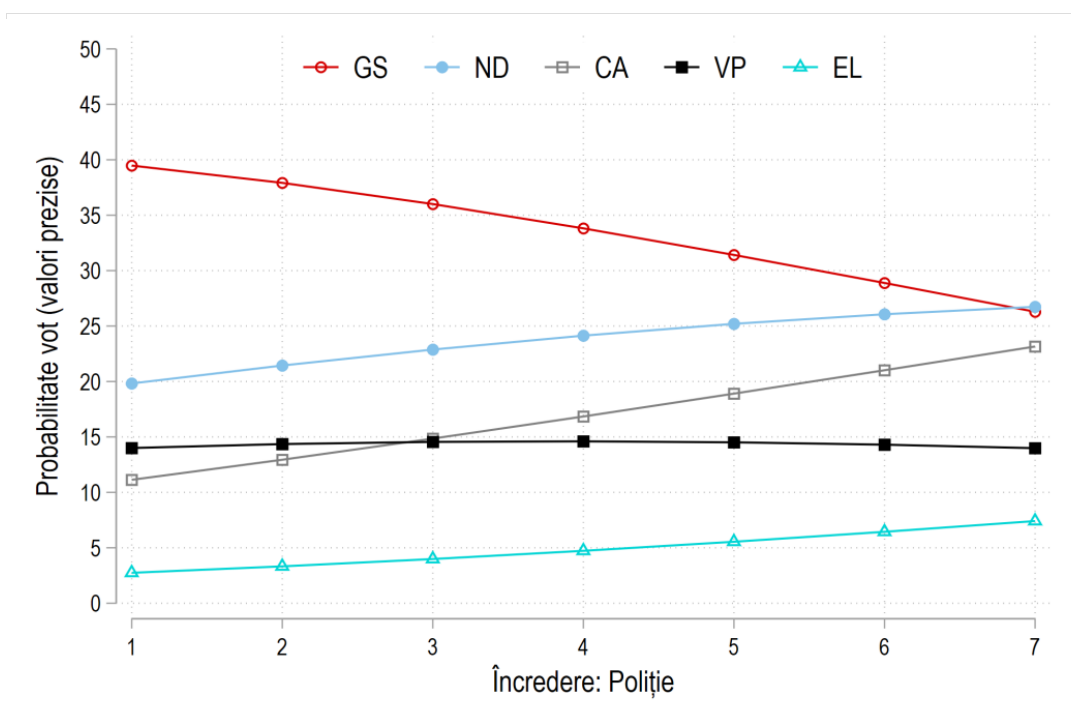
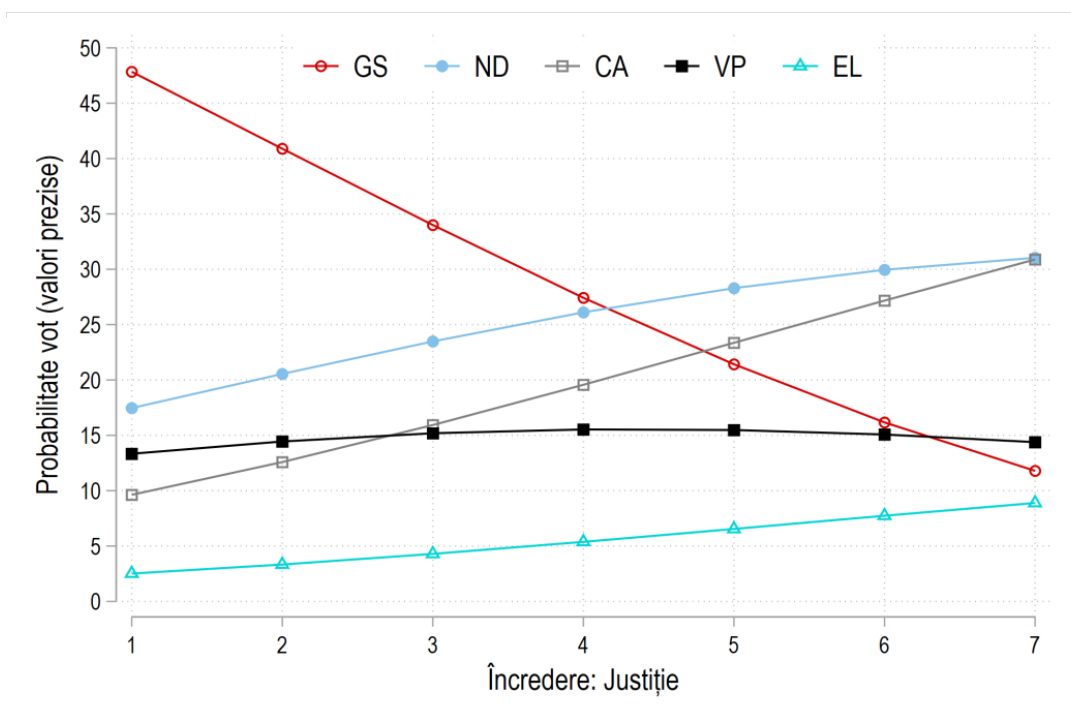
Probabilitatea de vot pentru Nicușor Dan variază destul de mult în funcție de încrederea în instituții. Observăm o relație pozitivă în cazul încrederii în Justiție, Poliție, Sistem Sanitar, Companii private, Presă și BNR: o încredere mai mare în aceste instituții este însoțită de o probabilitate mai mare de vot pentru Nicușor Dan. O relație negativă apare în cazul încrederii în Biserică.

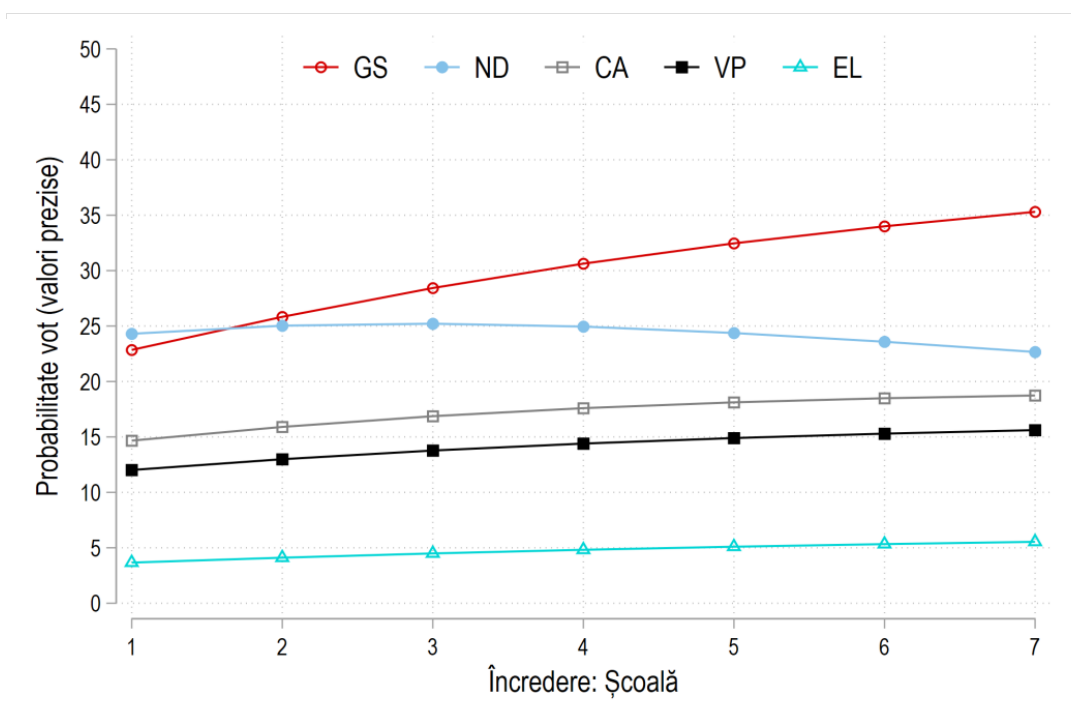
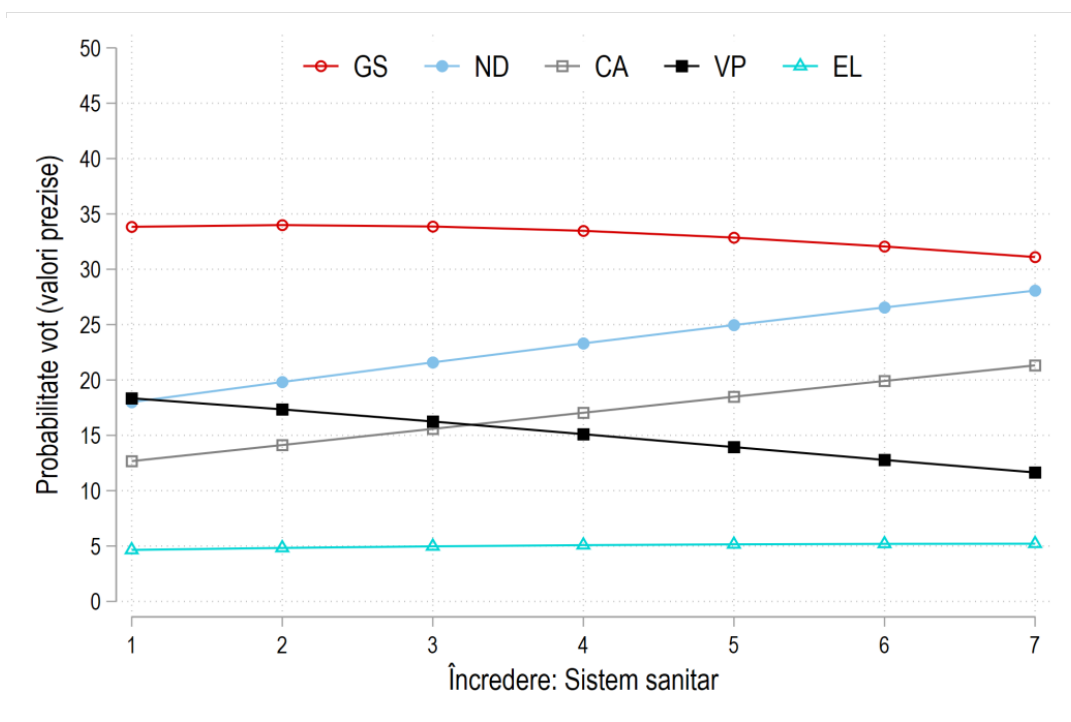
Probabilitatea de vot pentru George Simion variază destul de mult în funcție de încrederea în instituții. Comparativ cu Crin Antonescu, relația dintre vot și încredere este în general inversată. În două instanțe (Sistemul sanitar și Administrația publică locală) relația este redusă și ușor negativă. În cazul Bisericii este puternic pozitivă, iar în cazul Școlii pozitivă. În restul situațiilor relația este cel puțin negativă. O relație puternic negativă apare mai ales în cazul încrederii în Guvern, Partide, Justiție, Poliție, Companii private, Presă, BNR și alți oameni: o încredere scăzută în aceste instanțe este însoțită de o probabilitate mai mare de vot pentru George Simion.

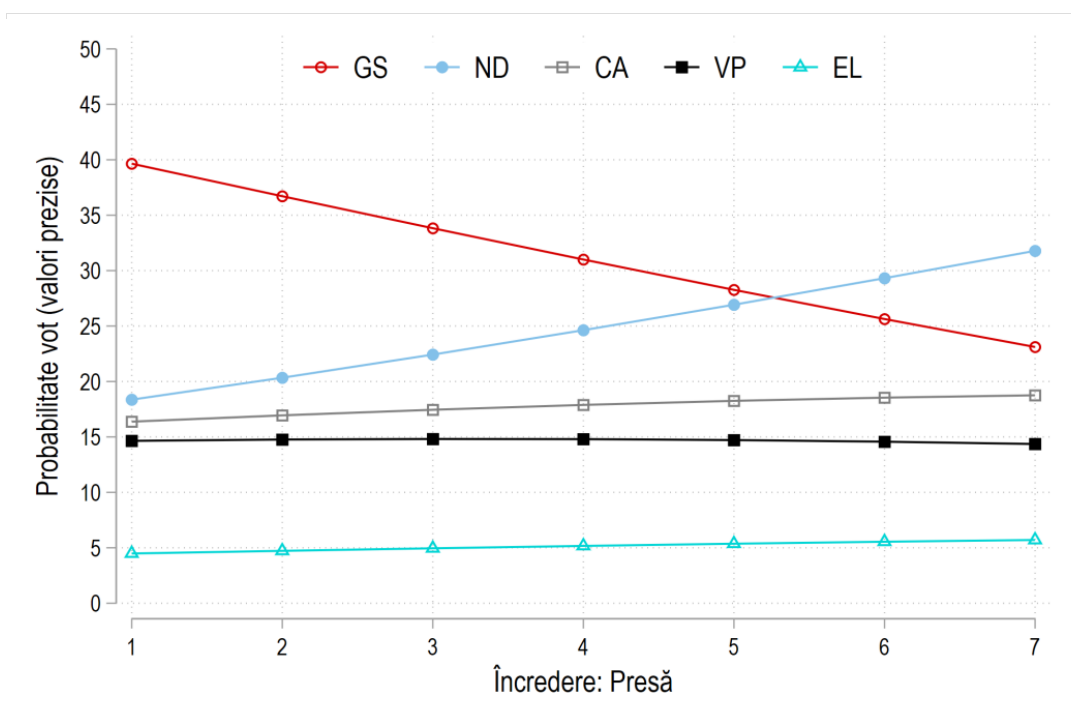
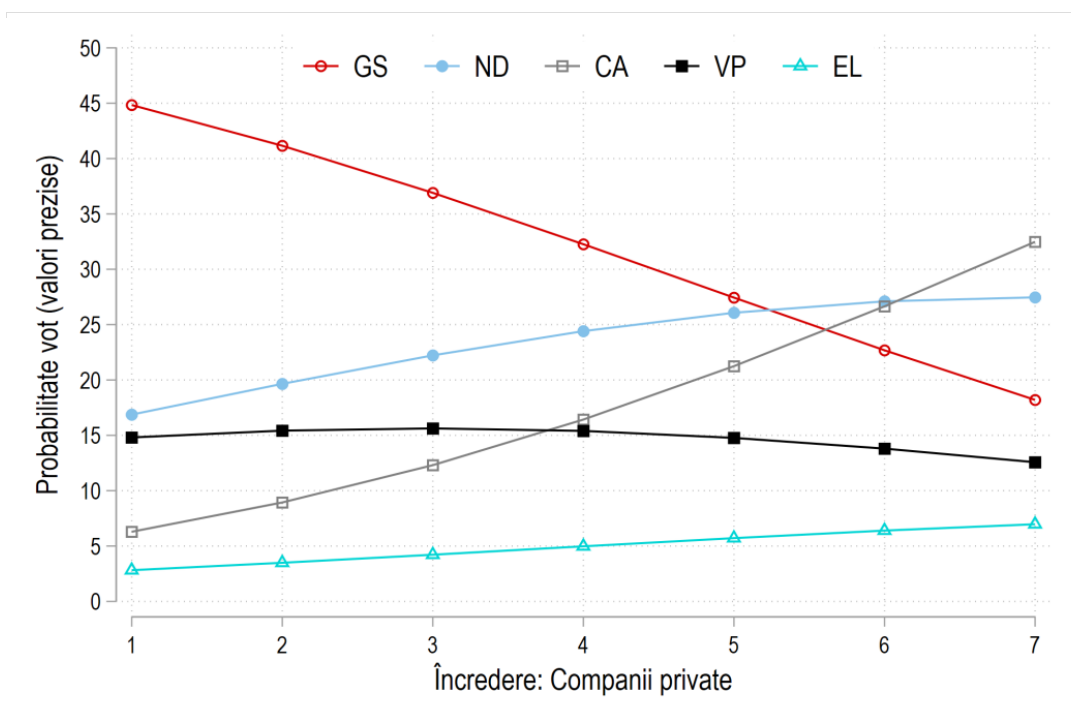
Pe baza acestor date, considerând toți candidații, pare că **opțiunea de vot este asociată cel mai puternic cu încrederea în trei instituții: Guvern, Partide și Justiție**. Un alegător fără încredere în aceste instituții are o probabilitate foarte mare de a vota cu George Simion. Un alegător cu încredere mare în aceste instituții are o probabilitate foarte mare de a vota cu Crin Antonescu. Restul alegătorilor se situează undeva între aceste extreme.

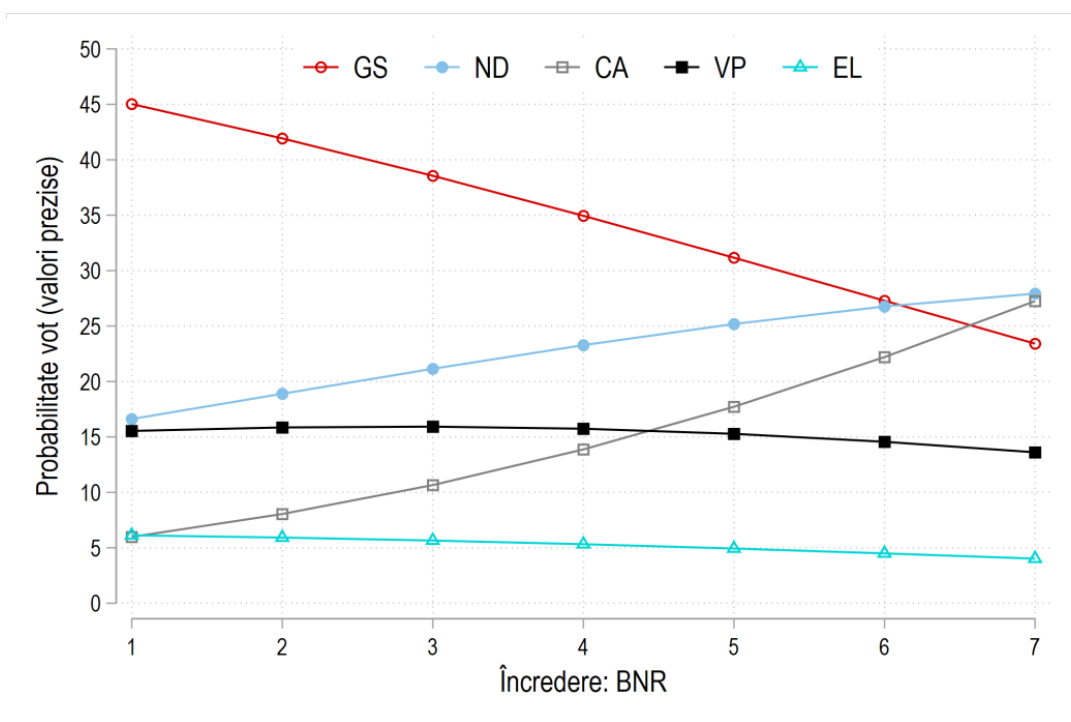
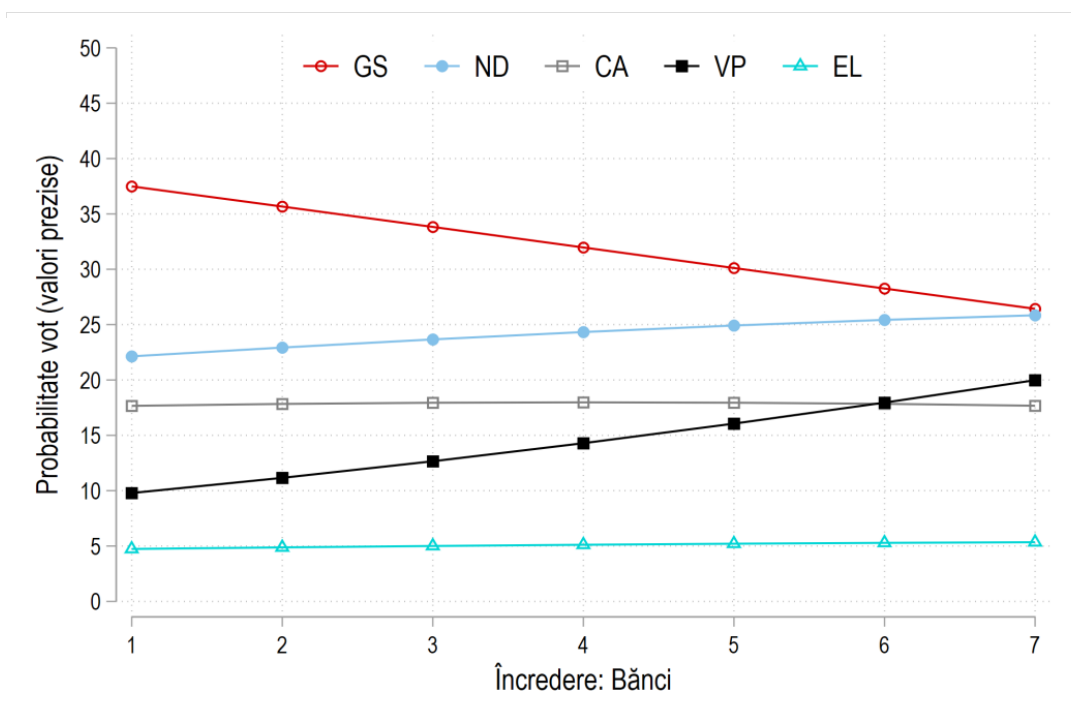


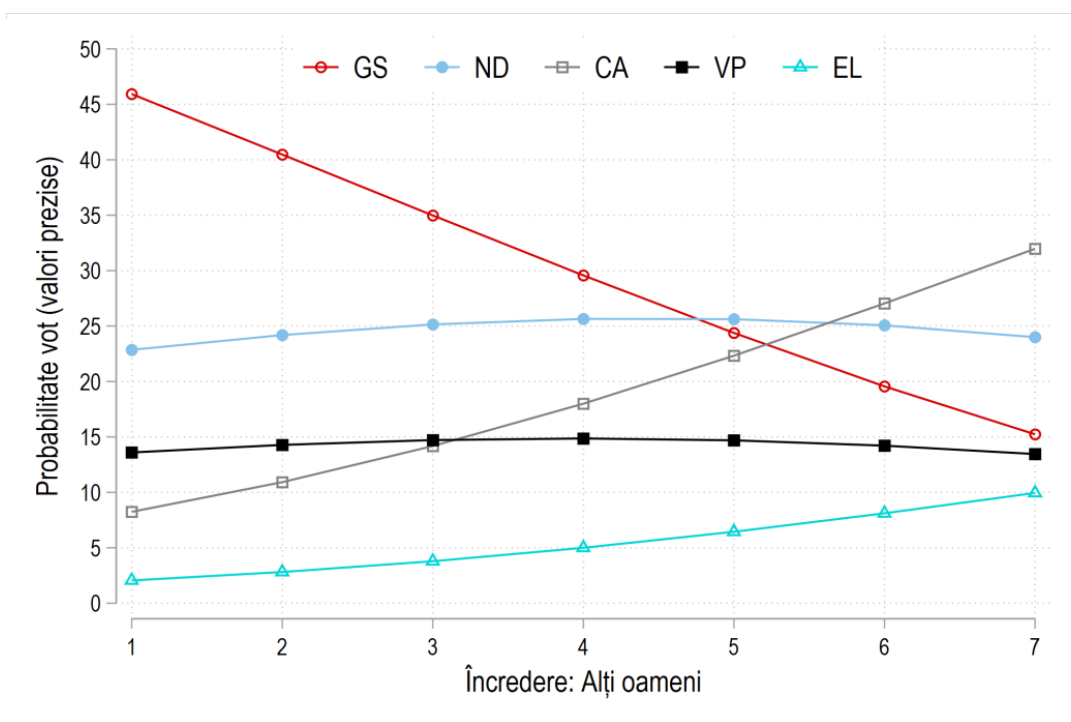
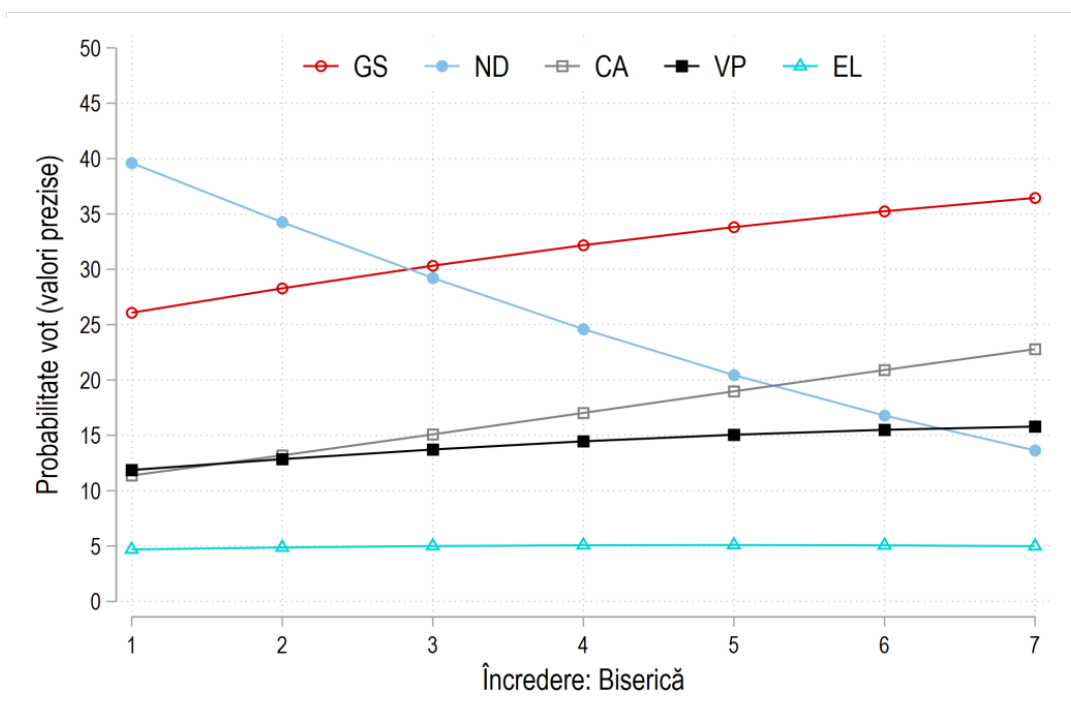












Dinamica intențiilor de vot: turul 1 al alegerilor prezidențiale din mai 2025 (2)

Data: 2025-04-15

Cuvinte cheie: alegerile-prezidentiale-2025, dinamica-intențiilor-de-vot

Cu mai bine de două săptămâni în urmă am făcut prima analiză pe tema dinamicii intențiilor de vot (vezi [aici](#)). Majoritatea precizărilor din postarea respectivă se aplică și aici, deci nu le mai reiau. Am făcut totuși unele modificări importante: (1) iau în calcul și estimările din sondajele apărute între timp, (2) prezint rezultatele a trei variante de modelare și (3) exclud două estimări în care lista candidaților prezentată respondenților a inclus-o și pe Diana Șoșoacă. Foarte important: faptul că nu am exclus din analiză niciun sondaj public – indiferent cine l-a realizat și cât de diferite au fost estimările respective de estimările din alte sondaje – crește incertitudinea, respectiv, foarte probabil, reduce calitatea estimărilor agregate. Astfel, includerea în analiză a ultimelor sondaje publicate (Sociopol, CURS, DataEcho, FlashData și Noi, cetățenii), unele cu estimări foarte diferite, chiar a produs astfel de efecte. Fără aceste sondaje, imaginea descrisă mai jos ar fi fost parțial diferită: George Simion ar fi avut o intenție de vot estimată mai mare, iar Nicușor Dan și Crin Antonescu mai mică.

Pentru prima parte a anului există un număr redus de sondaje, realizate mai degrabă de același institut (Sociopol). Ulterior, apar mai multe sondaje, realizate de mai multe institute, folosind metodologii diferite. Mult mai important, în unele cazuri, apar

diferențe foarte mari între estimările produse de sondaje. Pentru a corecta parțial aceste probleme, am testat trei modele statistice care ar putea sta în spatele estimărilor produse de sondaje. Rezultatul unei astfel de analize depinde în principal de modul în care modelăm relația dintre intenția de vot și numărul de zile până la alegeri, respectiv de diferențele dintre institutele de sondare („house effects”). Cele trei grafice prezentate mai jos diferă prin modul în care am modelat relația dintre timp și intenția declarată de vot: liniar, non-liniar și non-liniar GAM (Generalized Additive Models). Modelarea non-liniară e de fapt o medie ponderată între liniar și non-liniar (funcție polinomială de grad 2 aici). Date fiind diferențele dintre estimările produse de sondaje, estimarea de tip liniar este probabil mai apropiată de dinamica generală reală. Însă, probabil, tendințele recente sunt surprinse mai corect de modelele non-liniare.

Valorile prezentate în grafice nu reprezintă predicții ale procentelor pe care le au cu adevărat candidații sau pe care le vor obține, ci o imagine care descrie sintetic, în baza unor asumptii, estimările din sondajele publice. Simplu spus, în loc să ne uităm la estimările din toate sondajele, ne putem uita la tendințele din graficele prezentate aici.

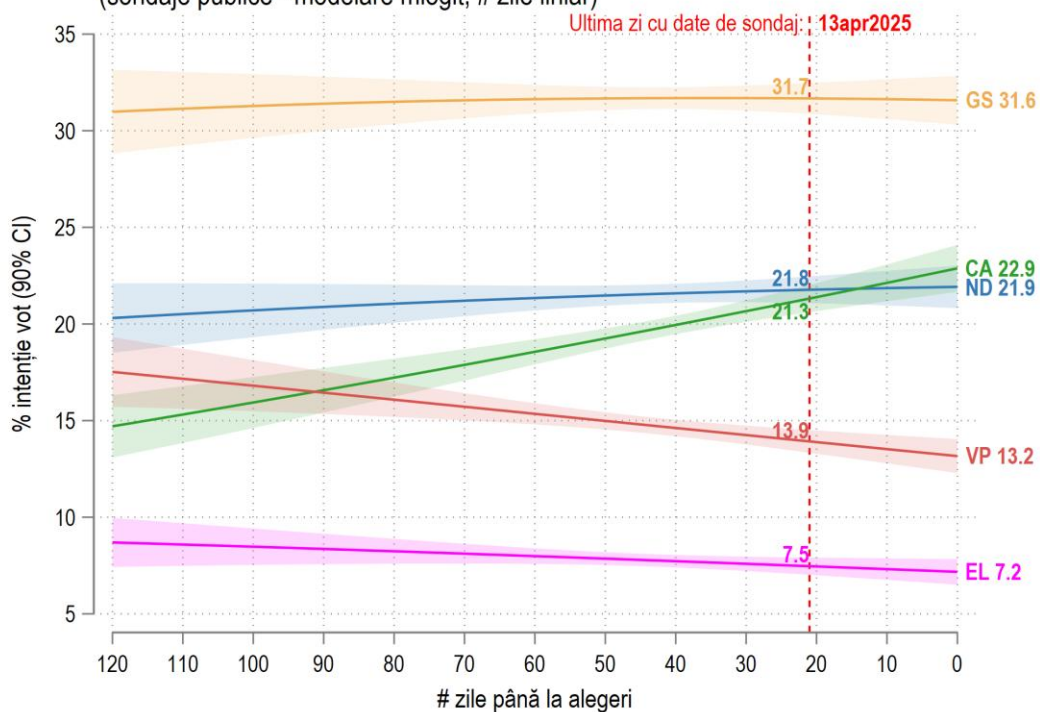
În comentariile scurte de mai jos prezint doar tendințele. Pentru a lua în calcul și incertitudinea estimărilor, ne uităm la ariile colorate asociate fiecărei estimări. Considerând strict estimările din sondajele publice, modelarea propusă a intențiilor de vot conform sondajelor publicate arată astfel:

- La data ultimului sondaj (13 aprilie), **George Simion** se situa pe prima poziție, cu o intenție de vot de 30-32%. La alegeri va obține probabil 32-33%. Funcție de model, trendul e ușor ascendent sau stabil.
- Pe poziția secundă, la distanță de locul 1, se situa probabil **Crin Antonescu** (diferența față de Nicușor Dan e mai mică decât marja de eroare), cu o intenție de vot de aproximativ 22-24%. La alegeri va obține probabil 23% sau chiar mai mult, dat fiind trendul ascendent.

- La data ultimului sondaj (13 aprilie), ultimul pe podium, se situa probabil **Nicușor Dan**, cu o intenție de vot de 22-24%. La alegeri va obține probabil 20-22%. Funcție de model, trendul e ușor descendent sau stabil. Sondajul DataEcho îl situează pe Nicușor Dan pe primul loc.
- Pe locul patru se situa **Victor Ponta**, cu o intenție de vot de 13-14%. La alegeri va obține probabil la fel. Funcție de model, trendul e mai degrabă ușor descendent sau stabil. Foarte important, sondajele prezintă estimări foarte diferite ale intențiilor de vot în cazul lui Victor Ponta. Astfel, sondajele realizate de Sociopol și cele comandate de USR îl poziționează în zona locurilor doi-trei, cu aproximativ 20% intenție de vot, în timp ce aproape toate sondajele online (AtlasIntel, FlashData, DataEcho și Noi, cetățenii) îl văd pe locul patru, cu cel mult 10% intenție de vot.
- Pe poziția cinci se situa **Elena Lasconi**, cu o intenție de vot de 7-8%. La alegeri va obține probabil 6-7%. Trendul e ușor descendent.

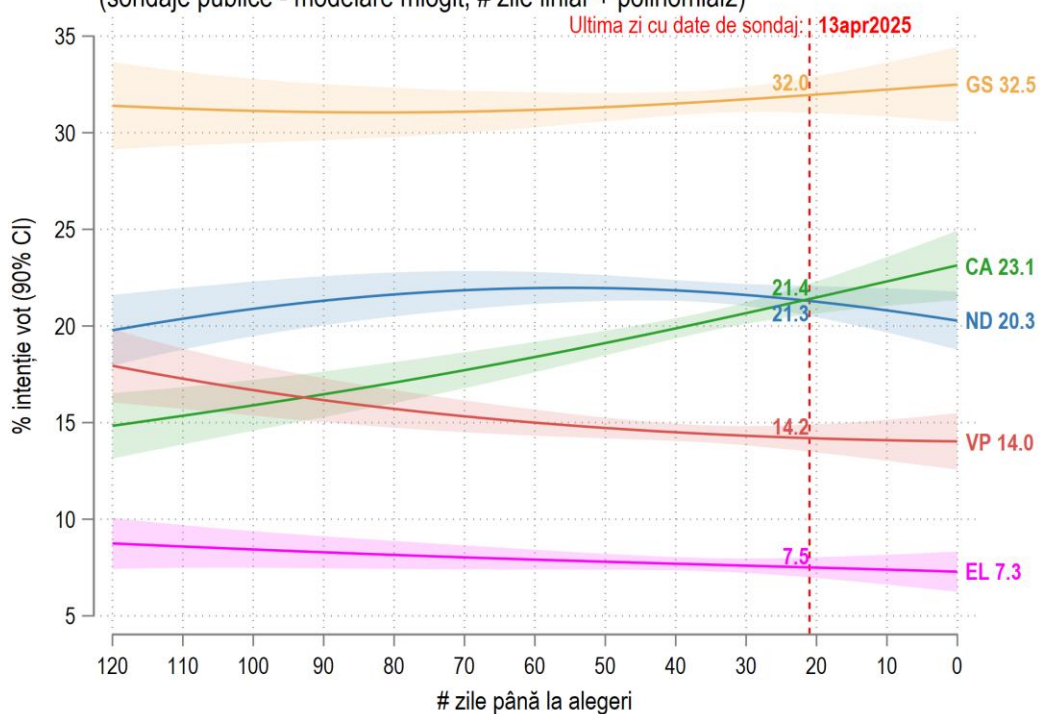
Prezidențiale 2025 (T1): Dinamica intențiilor de vot

(sondaje publice - modelare mlogit, # zile liniar)



Variabile de control: Institut & Tehnică anchetă (Online vs CATI)

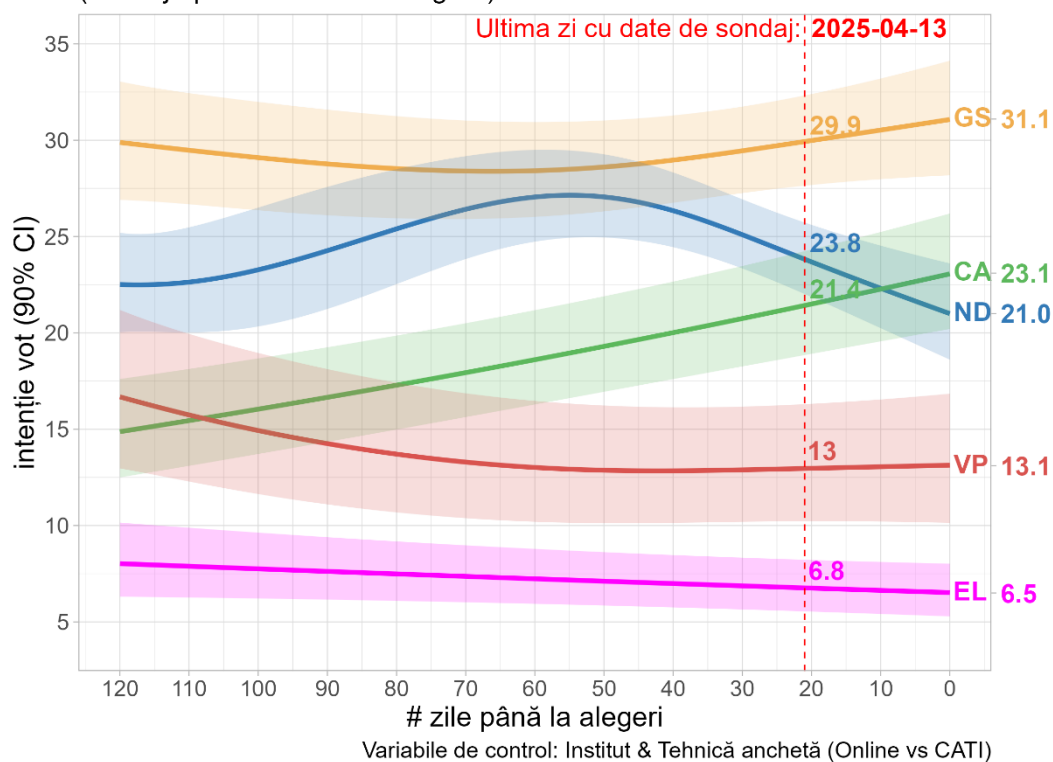
Prezidențiale 2025 (T1): Dinamica intențiilor de vot (sondaje publice - modelare mlogit, # zile liniar + polinomial2)



Variable de control: Institut & Tehnică anchetă (Online vs CATI)

Prezidențiale 2025 (T1): Dinamica intențiilor de vot

(sondaje publice - modelare gam)



„House effects”: alegerile prezidențiale 2025, T1 (2)

Data: 2025-04-16

Cuvinte cheie: alegerile-prezidentiale-2025, house-effects

În această postare reiau analiza pe tema „efectelor de institut” considerând toate sondajele publice apărute până în prezent (vezi analiza anterioară [aici](#)). De această dată prezint „efectul de institut” ca diferență între estimarea produsă de un institut de sondare și media estimărilor produse de toate institutele. O astfel de prezentare permite o comparare pe mai multe dimensiuni a efectelor. Desigur, ideal ar fi fost să calculez „efectul de institut” prin raportare la rezultatul final ... deocamdată necunoscut. Pentru că sondajele publicate au fost realizate în perioade diferite și au folosit metodologii diferite, o comparare directă a estimărilor este problematică. Pentru a crește comparabilitatea, folosesc și aici, similar cu analizele anterioare, predicții ale estimărilor. Toate predicțiile sunt realizate folosind un model GAM multivariat în care variabila independentă de interes este institutul. Modelul controlează influența perioadei în care au fost culese datele și tehnica de ancheta (CATI vs. Online).

Graficele de mai jos prezintă comparativ diferența dintre estimarea produsă de fiecare institut și media estimărilor institutelor (conform predicțiilor făcute de modelul statistic). Cercurile negre pline reprezintă estimările asociate candidaților, iar liniile portocalii intervalul de încredere al acestor estimări. Pentru a ușura lectura, am sortat estimările de la cea mai mare diferență negativă la cea mai mare diferență pozitivă. Dreptunghiul gri, denumit CI MEDIE, indică zona de incertitudine

asociată mediei estimărilor tuturor institutelor (intervalul de încredere / variație, cu o probabilitate de 90%). Valorile care se situează la stânga acestei zone sunt mai mici decât media, iar cele care se situează la dreapta sunt mai mari. Pentru a fi (aproape) sigur că un institut a produs o estimare semnificativ mai mică / mare comparativ cu media institutelor e nevoie ca întreg intervalul de variație al estimării (linia portocalie) să fie în afara zonei gri. Similar, putem compara estimările produse de două institute prin intermediul (in)existenței unei suprapunerii între intervalele de variație (liniile portocalii) asociate unuia dintre candidați.

Faptul că o anumită estimare nu atinge zona gri nu înseamnă automat că este și greșită, că supra-estimează / sub-estimează scorul unui candidat. O astfel de evaluare poate fi făcută doar după ce sunt publicate rezultatele alegerilor. În acest moment putem spune doar că o estimare asociată unui institut este mai mică / mare comparativ cu estimarea medie a tuturor institutelor, respectiv că este mai mică / mare comparativ cu estimarea făcută de un alt institut.

George Simion obține scoruri mai mici în cazul institutelor DataEcho, ARA, CURS și FlashData, respectiv mai mari în cazul Verifield și Altul (sondaj USR realizat de un institut necunoscut). Sondajul finanțat de USR indică un scor mai mare cu aproape nouă puncte procentuale decât media celorlalte sondaje.

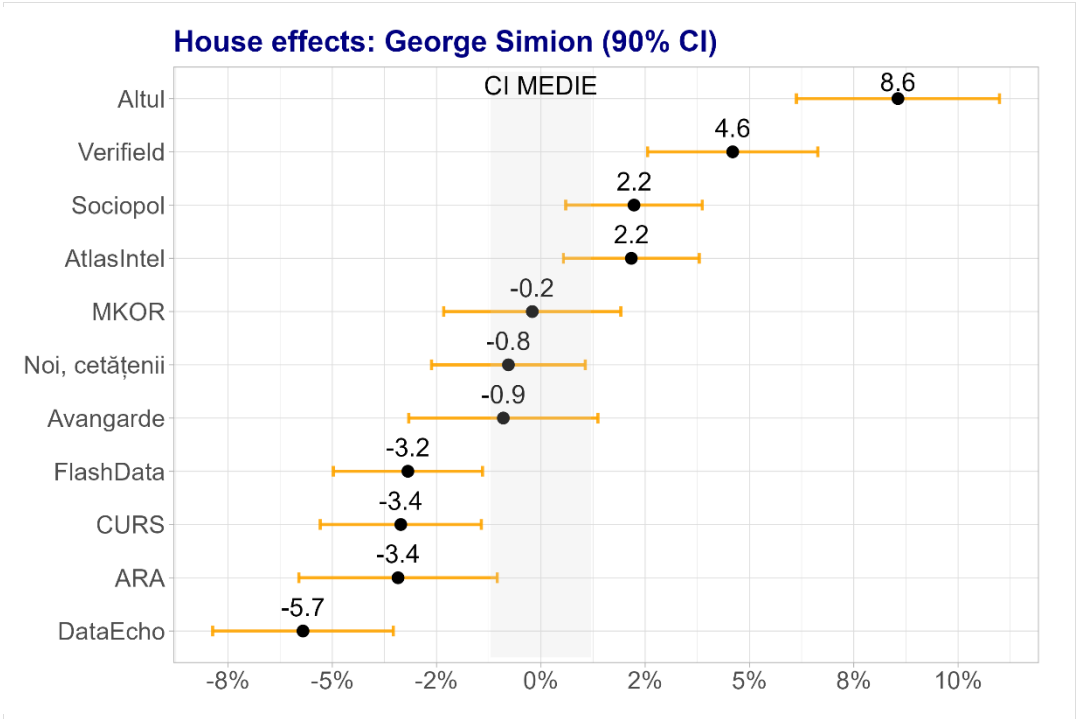
Nicușor Dan obține scoruri mai mici în cazul institutelor Sociopol, Altul, ARA și CURS (toate CATI), respectiv mai mari în cazul DataEcho, Noi, cetățenii și AtlasIntel (toate online). În cazul DataEcho și Noi, cetățenii diferența față de medie este de +6/7 puncte procentuale.

Crin Antonescu obține scoruri mai mici în cazul institutelor Sociopol, Altul și Verifield, respectiv mai mari în cazul FlashData, ARA, DataEcho și AtlasIntel. În cazul FlashData diferența față de medie este de +6 puncte procentuale.

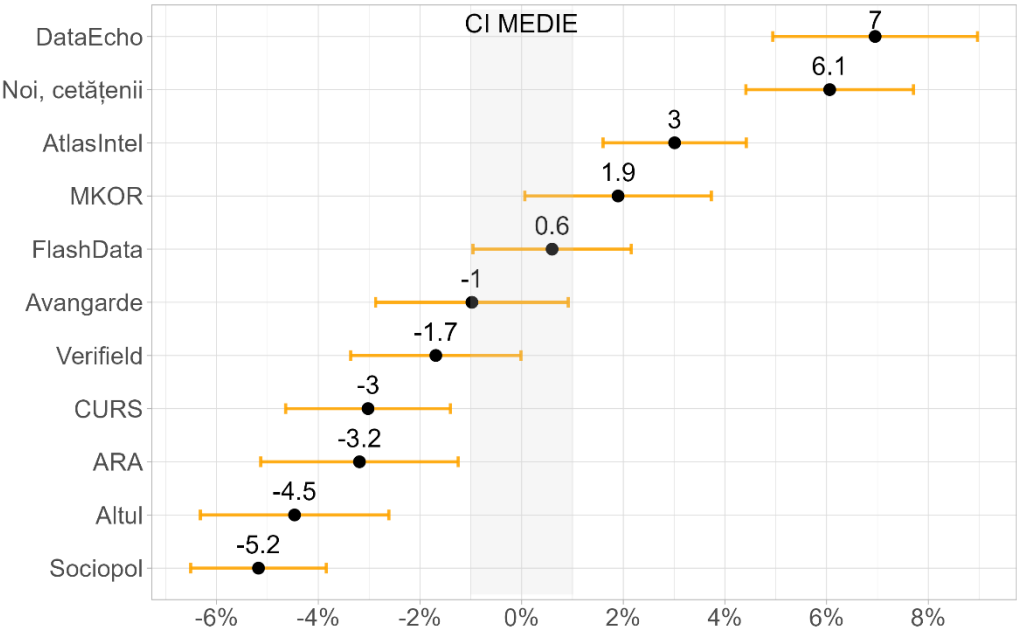
Victor Ponta obține scoruri mai mici în cazul institutelor FlashData, AtlasIntel, Noi, cetățenii, DataEcho și ARA (toate online, cu excepția ARA), respectiv mai mari în cazul Sociopol,

Verifield și Altul. În cazul celor trei institute „suplimentul” de scor este consistent (5-8 puncte procentuale).

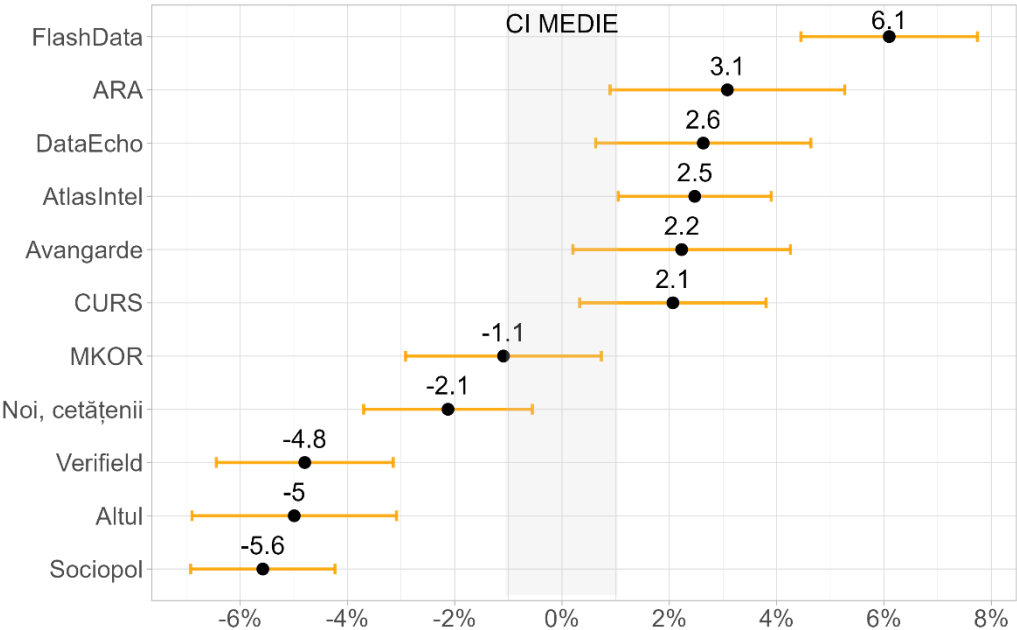
Elena Lasconi obține scoruri mai mici în cazul institutelor Altul (USR fără nume institut), AtlasIntel, MKOR și Verifield, respectiv mai mari în cazul ARA, Sociopol și CURS.



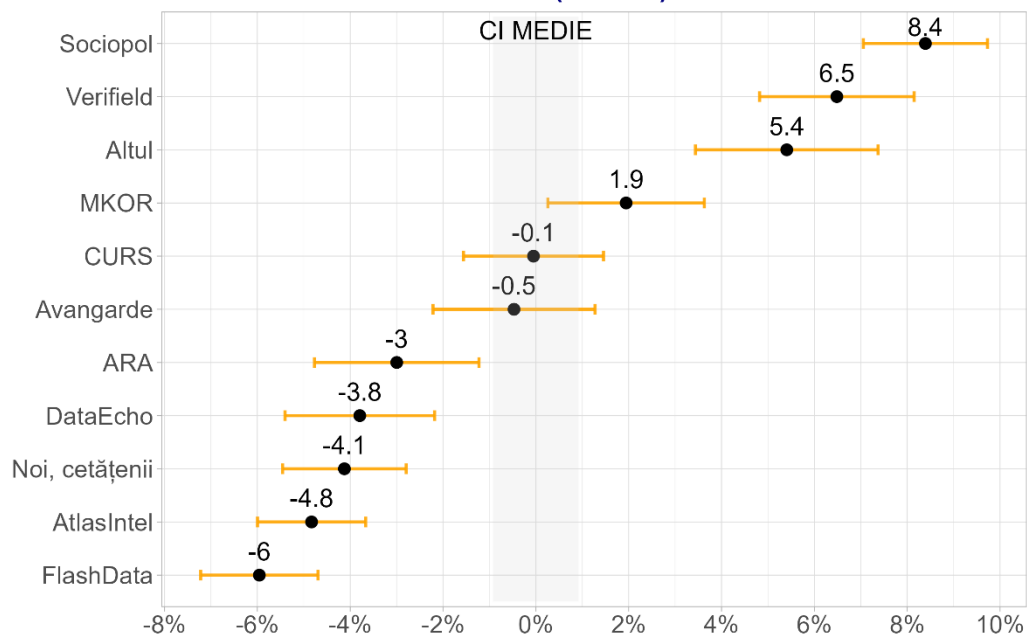
House effects: Nicușor Dan (90% CI)



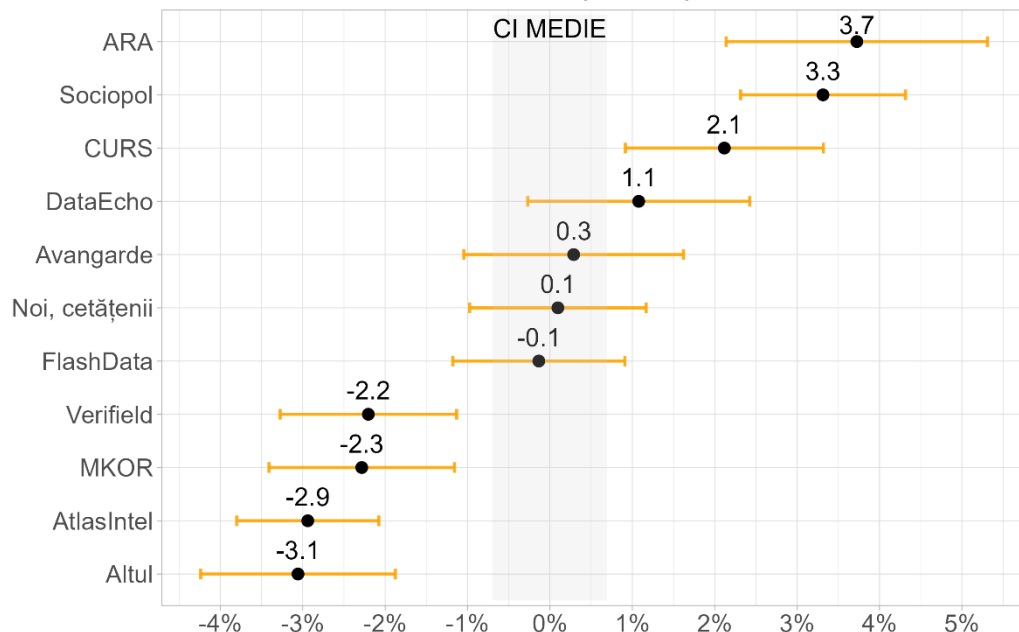
House effects: Crin Antonescu (90% CI)



House effects: Victor Ponta (90% CI)



House effects: Elena Lasconi (90% CI)



Efecte de metodă: CATI vs. Online

Data: 2025-04-17

Cuvinte cheie: alegerile-prezidentiale-2025, cati-vs-online

Scopul acestei analize este de a vedea dacă tehnica de sondaj (modalitatea de colectare a datelor) are un impact asupra estimărilor obținute. Pentru că un Institut folosește de obicei o singură tehnică, e dificil de atribuit sursa diferențelor unuia dintre cele două tipuri de efecte. Simplu spus, cu datele existente, nu putem estima cu certitudine care este cauza diferențelor dintre estimările observate ale intențiilor de vot: institutul, tehnica de sondaj sau, mai probabil, o combinație a acestora.

Problema este chiar mai complicată de atât. Alături de cele două cauze deja menționate, există o serie de alte cauze de natură metodologică: perioada de culegere a datelor, populația vizată (alegători vs. votanți probabili), procentul cazurilor valide, tratarea non-răspunsurilor, formula de calcul a ponderărilor (weights), formularea și ordinea întrebărilor, imaginea institutelor de sondare la nivelul opiniei publice etc. Pentru că informațiile despre modul în care au fost produse estimările din sondajele publice sunt limitate, efectul acestor cauze nu poate fi controlat. În aceste condiții există riscul să „descoperim” că principala cauză a diferențelor dintre estimările produse de sondaje are legătură cu modalitatea în care au fost colectate datele (CATI, online), în timp ce cauzele reale pot fi mai degrabă altele. O altă limită este dată de premisa de la care plecăm atunci când lucrăm cu aceste estimări: „nu s-a umblat la date,” adică institutele nu au distorsionat sau modificat intenționat

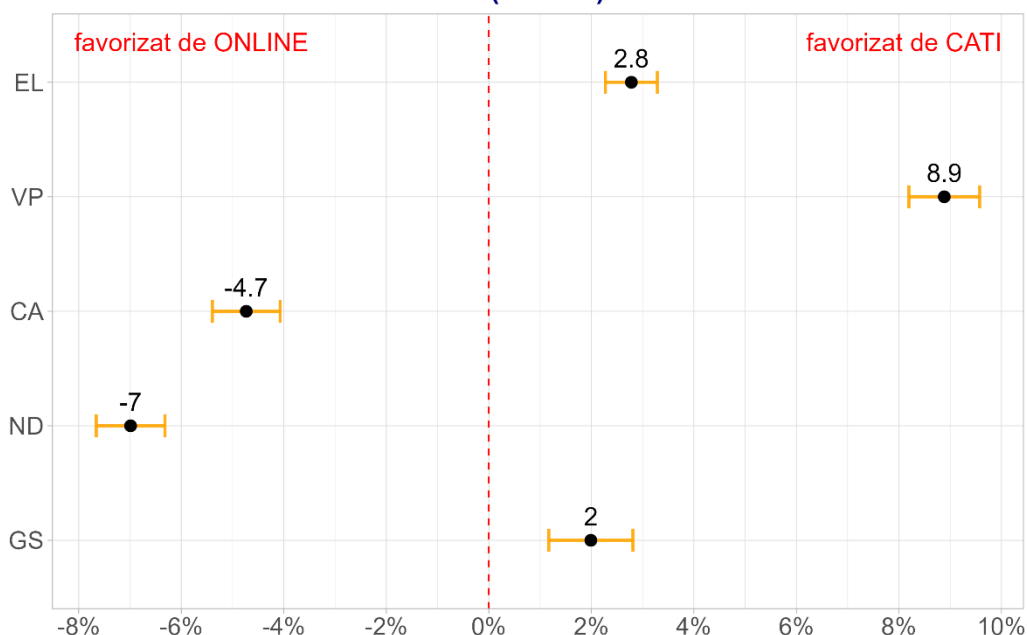
estimările. În concluzie, estimările prezentate aici sunt mai degrabă orientative.

Toate sondajele publice realizate în acest an au fost de tip CATI sau Online. CATI se referă la colectarea datelor prin intermediul telefonului (Random Digital Dialing, eșantion de numere selectate aleator sau alte variante), iar Online prin intermediul Internetului (recrutare digitală aleatorie, panel recrutat online sau alte variante). Pentru că informațiile lipsesc cel mai adesea, analizez doar diferența dintre CATI (telefonic) și online.

Strategia de analiză este aceeași: regresie logistică multivariată (GAM), cu variabila independentă de interes tehnica și variabilele de control data aplicării chestionarelor și institutul. Rezultatul analizei prezintă, sub forma de puncte procentuale, diferența dintre două predicții realizate pe baza datelor din sondaje: predicția CATI - predicția online. O valoare apropiată de zero indică absența unei diferențe (a unui efect de metodă), o valoare pozitivă indică o estimare mai mare a intenției de vot pentru un anumit candidat în sondajele CATI, iar o valoare negativă o estimare mai mare în sondajele online. Estimările obținute sunt prezentate ca valori (cercurile negre pline) și ca intervale de variație probabile (liniile portocalii).

Concluzii: Sondajele online au produs estimări mult mai mari în cazul lui Nicușor Dan (+7 puncte procentuale) și, oarecum surprinzător, mai mari în cazul lui Crin Antonescu (5 pp; mă așteptam ca scorul acestui candidat să fie ceva mai mare în sondajele CATI). Sondajele telefonice (CATI) produc estimări mult mai mari în cazul lui Victor Ponta (9 pp, probabil o supra-estimare) și puțin mai mari în cazul candidaților George Simion (2 pp) și Elena Lasconi (3 pp).

Efect de metodă: CATI - Online (90% CI)



Cum pot fi folosite aceste estimări? Atunci când vedem o estimare a intenției de vot pentru un anumit candidat, aflăm care a fost modalitatea de colectare a datelor și, funcție de situație, scădem sau adunăm din acea estimare numărul de puncte procentuale indicat în grafic. Obținem astfel estimarea intenției de vot care s-ar fi obținut dacă sondajul ar fi fost realizat prin cealaltă tehnică. Dacă dorim să ținem cont de ambele tehnici, facem media celor două estimări (alternativ scădem/ adunăm jumătate din numărul de puncte procentuale indicat în grafic). De exemplu, dacă într-un sondaj telefonic (CATI) candidatul Victor Ponta are o estimare a intenției de vot de 20%, într-un sondaj online, celelalte condiții fiind similare, are probabil aproximativ 11% ($20\% - 9\%$). Calculând media simplă dintre 20% și 11% obținem 15.5%, estimare care în general are șanse mai mari să fie mai apropiată de realitate (alternativ, scădem 5.5% din 20%). Desigur, aceste calcule corespund situației în care nivelul de încredere în cele două tehnici de sondare este aproximativ același.

În aceeași logică, dacă dorim să comparăm estimările din două sondaje realizate în aproximativ aceleași condiții, cu excepția

tehnicii de colectare a datelor, adunăm / scădem punctele procentuale din grafic la estimările realizate de unul dintre sondaje. De exemplu, să presupunem că Crin Antonescu a obținut într-un sondaj CATI o intenție de vot de 20%, iar într-un sondaj online 25%. Pentru că acest candidat obține în sondajele online cu aproximativ 5 pp mai mult decât în cele CATI, putem concluziona că cele două estimări sunt relativ similare. Desigur, putem evalua care tehnică a produs estimările cele mai corecte doar după aflarea rezultatelor oficiale.

Cine e pe locul 2?

Data: 2025-04-24

Cuvinte cheie: alegerile-prezidentiale-2025, cati-vs-online

Răspunsul la întrebarea din titlu diferă adesea în funcție de cine și cum a cules datele. Am semnalat în ultimele două postări faptul că estimările intențiilor de vot variază mult în funcție de institut și tehnica de colectare a datelor (telefonic vs. online). Graficul de azi ne arată cum variază estimările intențiilor de vot în funcție de două variabile: tehnica folosită (telefonic vs. online) și experiența institutului care a produs estimările (institut nou vs. vechi; am considerat că un institut e nou dacă nu a publicat sistematic estimări ale intențiilor de vot cu ocazia ultimelor două runde de alegeri).

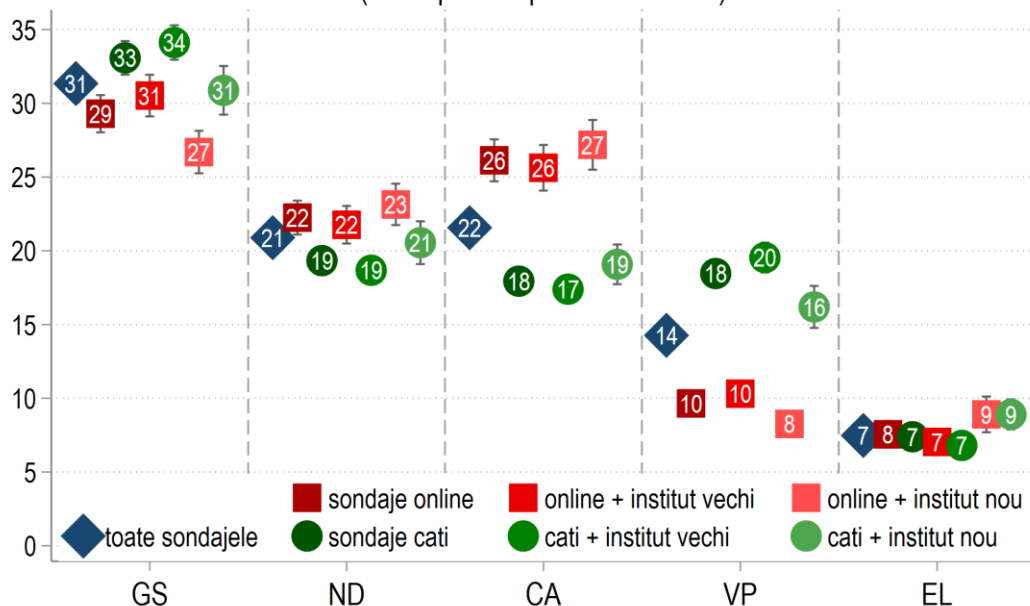
Valorile prezentate în grafic sunt predicții pe baza unui model de regresie multinominală în care am inclus două variabile de control: numărul de zile până la alegeri și o evaluare personală, (parțial) subiectivă, cu privire la existența unei posibile favorizări, nu neapărat intenționate, a unui candidat din partea unui institut. Am considerat că favorizarea are loc atunci când există un potențial conflict de interese între rolul de „sondor” și alte roluri (consultant campanie, servicii de colectare de date, membru sau simpatizant partid / candidat, istoric relevant comun). Despre favorizare voi discuta într-o analiză viitoare.

Valorile din grafic prezintă intențiile de vot prezise de model pentru ziua de 14 aprilie (20 de zile înainte de turul 1). Pentru fiecare candidat am estimat șapte valori: toate sondajele, sondajele online, sondajele CATI, sondajele online realizate de institute cu vechime, sondajele online realizate de institute noi,

sondajele CATI realizate de institute cu vechime, sondajele CATI realizate de institute noi. În cazul candidaților **Elena Lasconi și Nicușor Dan** variațiile intențiilor de vot sunt mici, deci, destul de probabil, estimările sunt apropiate de realitatea de la acest moment (14 aprilie). În cazul lui **George Simion** variațiile sunt puțin mai mari: intenția de vot pentru acest candidat este puțin mai mare în cazul sondajelor telefonice, respectiv a institutelor cu vechime (sondajele online realizate de institute noi produc estimarea cea mai mică). Intenția de vot pentru Crin Antonescu, respectiv Victor Ponta variază foarte mult: **Crin Antonescu** obține scoruri mai mari în cazul sondajelor online (surprinzător și ciudat, aș zice, dat fiind profilul probabil al electoratului acestui candidat), iar **Victor Ponta** obține scoruri mai mari în cazul sondajelor telefonice.

Pe baza estimărilor din grafic, **sunt șanse mari ca, la acest moment (14 aprilie), primul loc să fie ocupat de George Simion, iar ultimul de Elena Lasconi. Pe locul secund, funcție de datele la care ne uităm, se poate situa oricare dintre ceilalți trei candidați.** Pentru că (1) sondajele telefonice sunt, în general, mai potrivite pentru contextul din România (mai ales că populația de interes este cea națională) și (2) în mod normal institutele care activează de mai multă vreme pe piață ar trebui să producă estimări mai corecte, am putea aprecia că **intenția de vot pentru fiecare dintre cei trei candidați se situează undeva în intervalul 16-21%, prin urmare lupta pentru locul doi este foarte strânsă.** Dacă alegem să avem încredere ceva mai mare în sondajele online, respectiv în institutele nou apărute, **lupta pentru locul secund se dă doar între Crin Antonescu și Nicușor Dan, primul având probabil un ușor avans.**

Intenția de vot în funcție de tehnica de sondaj și vechimea institutului
(valori prezise pentru ziua T-20)



Variația estimărilor intențiilor de vot în funcție de partizanatul potențial al institutelor de sondare

Data: 2025-04-25

Cuvinte cheie: alegerile-prezidentiale-2025, house-effects

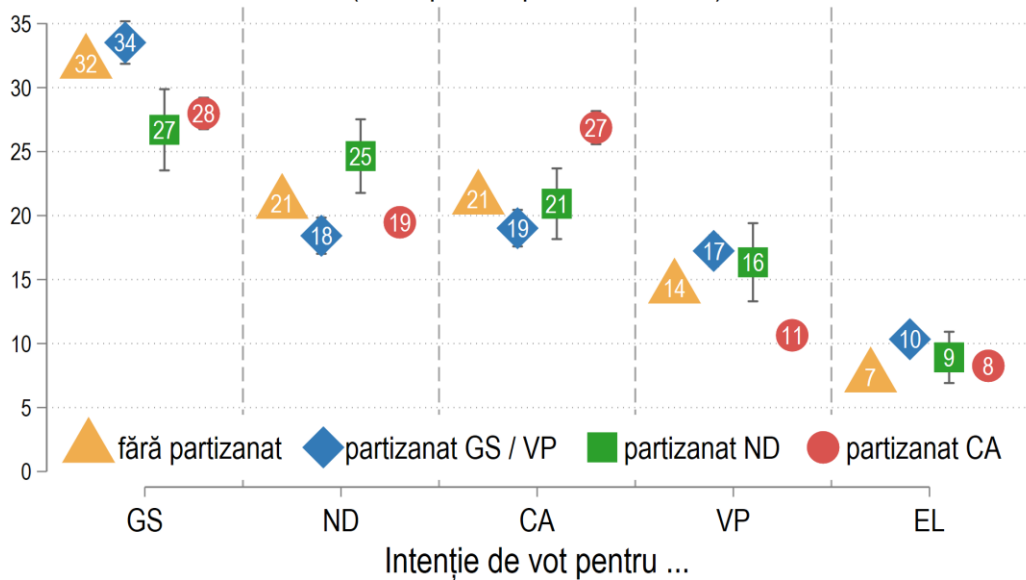
Uneori, institutelor de sondare a opiniei publice produc, cu sau fără intenție, estimări distorsionate ale preferințelor de vot. Fără informații de culise e dificil de spus cu certitudine care institute sunt partizane în relație cu care dintre candidați. **Pornind de la informațiile prezentate în media, am evaluat, (parțial) subiectiv, partizanatul potențial al tuturor institutelor care au publicat estimări ale intențiilor de vot. Am considerat că partizanatul / favorizarea are loc atunci când există un potențial conflict de interese între rolul de „sondor” și alte roluri precum consultant în campania electorală, furnizarea unor servicii de colectare de date, calitatea de membru sau simpatizant al unui partid / candidat, relații anterioare relevante.**

Valorile prezentate în grafic sunt predicții pe baza unui model de regresie multinominală: variabila dependentă este intenția de vot, independenta de interes este evaluarea partizanatului institutelor, iar variabilele de control sunt numărul de zile până la alegeri, modul de colectare a datelor (telefonice vs. online) și experiența institutului care a produs estimările (institut nou vs. vechi; am considerat că un institut e nou dacă nu a publicat sistematic estimări ale intențiilor de vot cu ocazia ultimelor două runde de alegeri).

Valorile din grafic prezintă intențiile de vot prezise de model pentru ziua de 14 aprilie (20 de zile înainte de turul 1). Pentru fiecare candidat am estimat patru valori în funcție de partizanatul potențial al institutelor pentru un anumit candidat: fără partizanat, partizanat în favoarea lui George Simion /Victor Ponta (cei doi candidați par să fie favorizați de același institut și doar de acesta, prin urmare nu am putut distinge între cele două posibile efecte), partizanat pentru Nicușor Dan și partizanat pentru Crin Antonescu. În cazul candidatei Elena Lasconi nu am avut informații pentru a o asocia cu un anumit institut. În cazul candidatului Nicușor Dan estimările au un grad de incertitudine mare pentru că am identificat un singur institut posibil partizan, iar acesta a realizat un singur sondaj.

Datele din grafic arată clar că, **atunci când există un partizanat al unui institut în favoarea unui candidat, candidatul respectiv obține un scor mai mare în sondajele realizate de acel institut și, firesc, contra-candidații obțin un scor mai mic.** Astfel, ținând cont de sondajele realizate de institutele nepartizane, intenția de vot pentru George Simion este 32%; scorul lui Simion crește la 34% în sondajele realizate de un institut care îi este favorabil și scade la 27-28% în sondajele realizate de un institut favorabil altor candidați. În cazul lui Nicușor Dan estimarea nepartizană este 21%, cea partizană 25%, iar cea partizană în relație cu alți candidați 18-19%. În cazul lui Crin Antonescu valorile sunt, în aceeași ordine, 21%, 27% și 19-21%. În cazul lui Victor Ponta valorile sunt 14%, 17% și 11-16%. În concluzie, **comparativ cu institutele care sunt partizane în relație cu alți candidați, institutele partizane obțin pentru candidatul preferat scoruri mai mari cu aproximativ 6-8 puncte procentuale.** Dacă avem în vedere doar estimările institutelor nepartizane și a celor partizane în relație cu candidatul preferat, diferențele sunt mai mici: George Simion obține în estimările institutelor partizane un scor mai mare cu două puncte procentuale, Nicușor Dan patru puncte, Crin Antonescu șase puncte, iar Victor Ponta trei puncte.

Intențiile de vot în funcție de partizanatul potențial al institutelor (valori prezise pentru ziua T-20)



Dinamica intențiilor de vot: turul 1 al alegerilor prezidențiale din mai 2025 (3)

Data: 2025-04-28

Cuvinte cheie: alegerile-prezidentiale-2025, dinamica-intențiilor-de-vot, polls-of-polls

În această postare reiau analizele anterioare pe tema dinamicii intențiilor de vot (vezi [aici](#) și [aici](#)). Majoritatea precizărilor din postările respective se aplică și în cazul acestei analize, deci nu le mai reiau. Desigur, am adăugat estimările produse de sondajele nou apărute. Suplimentar, am adăugat un tip nou de analiză (bayesian state-space model).

Valorile prezentate în grafice nu reprezintă predicții ale procentelor pe care le au cu adevărat candidații sau pe care le vor obține, ci o imagine care descrie sintetic, în baza unor asumții, estimările din sondajele publicate până la data de 23 aprilie. Simplu spus, în loc să ne uităm la estimările din toate sondajele, ne putem uita la tendințele din graficele prezentate aici.

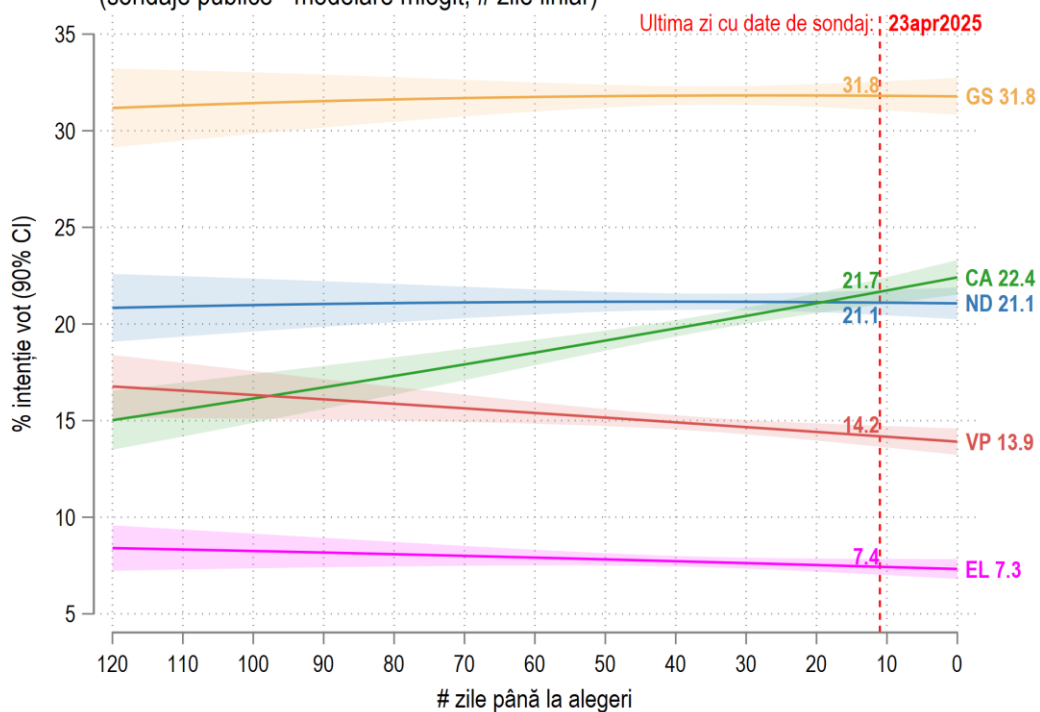
În comentariile de mai jos prezint doar tendințele observate în grafice. Pentru a lua în calcul și incertitudinea estimărilor, ne uităm la ariile colorate asociate fiecărei estimări. Considerând strict estimările din sondajele publice, modelarea propusă a intențiilor de vot conform sondajelor publicate arată astfel:

- La data ultimului sondaj (23 aprilie), **George Simion** se situa pe prima poziție, cu o intenție de vot de aproximativ 32%. La alegeri va obține probabil un scor

similar (în țară; cu voturile din diaspora, va ajunge la aproximativ 34%). Trendul e în general stabil.

- Pe poziția secundă, la distanță de locul 1, se situa probabil **Crin Antonescu** (diferența față de Nicușor Dan e mai mică decât marja de eroare), cu o intenție de vot de aproximativ 21-22%. Dat fiind trendul ascendent, la alegeri va obține probabil 23% sau ceva mai mult.
- La data ultimului sondaj (23 aprilie), ultimul pe podium, se situa probabil **Nicușor Dan**, cu o intenție de vot de 19-21%. La alegeri va obține probabil 19-21%. Funcție de model, trendul e ușor descendent sau stabil. Sondajul DataEcho îl situează pe Nicușor Dan pe primul loc.
- Pe locul patru se situa **Victor Ponta**, cu o intenție de vot de 14-15%. La alegeri va obține probabil tot 14-15%, funcție și de gradul real de disciplinare a electoratului PSD. Trendul depinde foarte mult de model. Foarte important, sondajele prezintă estimări foarte diferite ale intențiilor de vot în cazul lui Victor Ponta. Astfel, sondajele realizate de Sociopol și cele comandate de USR îl poziționează în zona locurilor doi-trei, cu aproximativ 20% intenție de vot, în timp ce aproape toate sondajele online (AtlasIntel, FlashData, DataEcho și Noi, cetățenii) îl văd pe locul patru, cu cel mult 10% intenție de vot.
- Pe poziția cinci se situa **Elena Lasconi**, cu o intenție de vot de 7-8%. La alegeri va obține probabil 7-8%. Trendul diferă ușor în funcție de model, dar tinde să fie stabil.

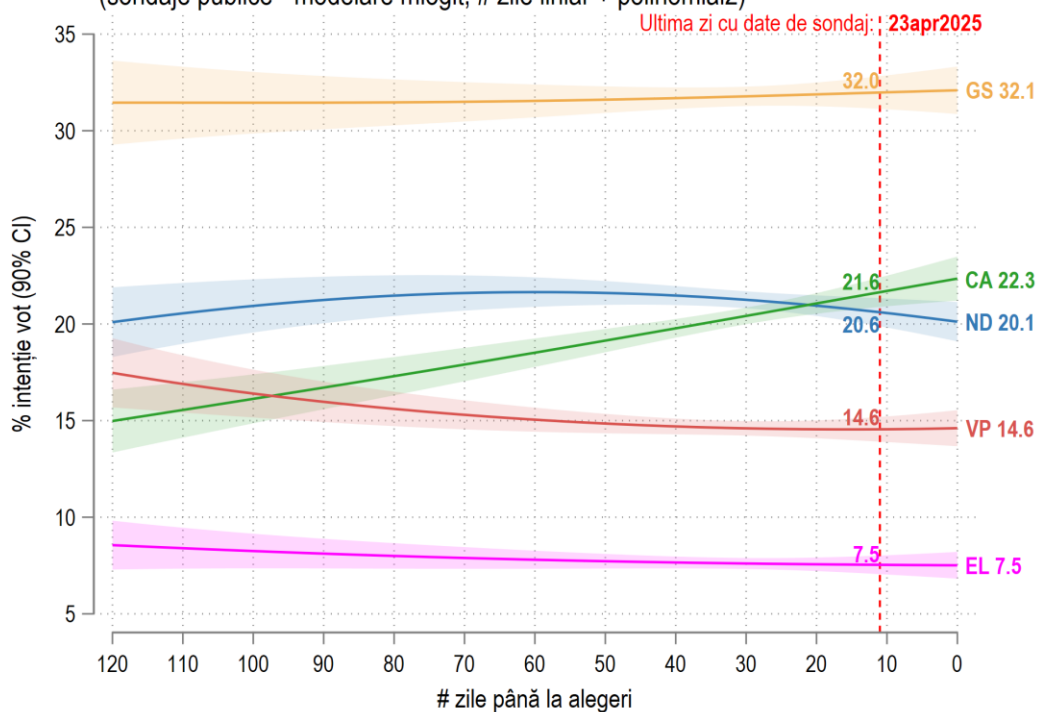
Prezidențiale 2025 (T1): Dinamica intențiilor de vot (sondaje publice - modelare mlogit, # zile liniar)



Variable de control: Institut & Tehnică anchetă (Online vs CATI)

Prezidențiale 2025 (T1): Dinamica intențiilor de vot

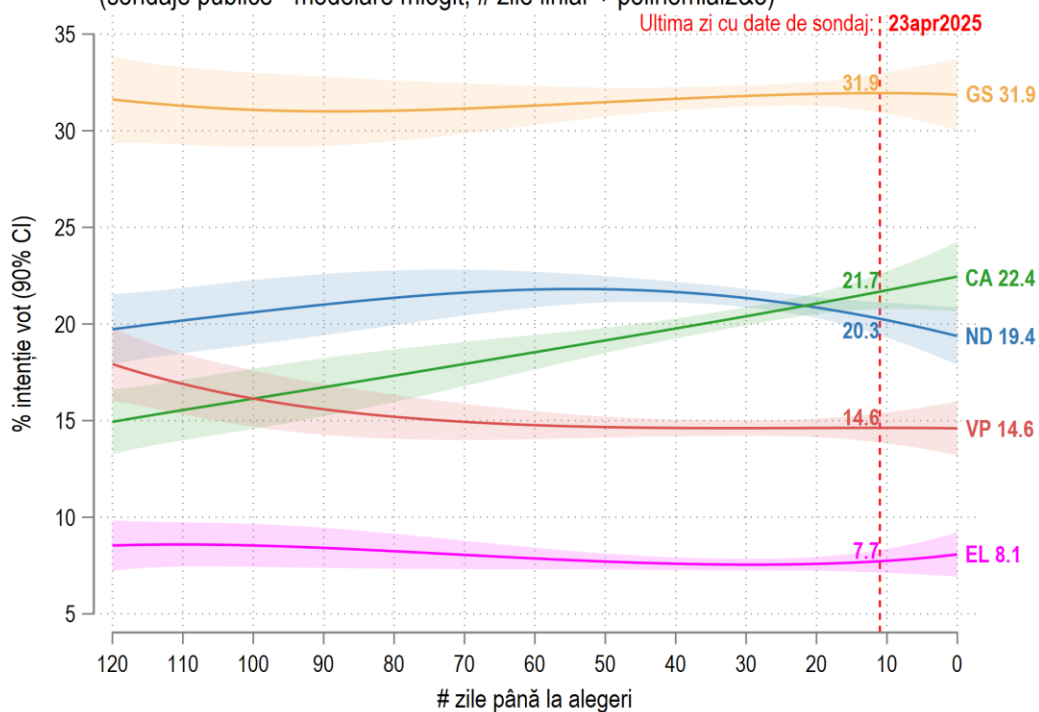
(sondaje publice - modelare mlogit, # zile liniar + polinomial2)



Variabile de control: Institut & Tehnică anchetă (Online vs CATI)

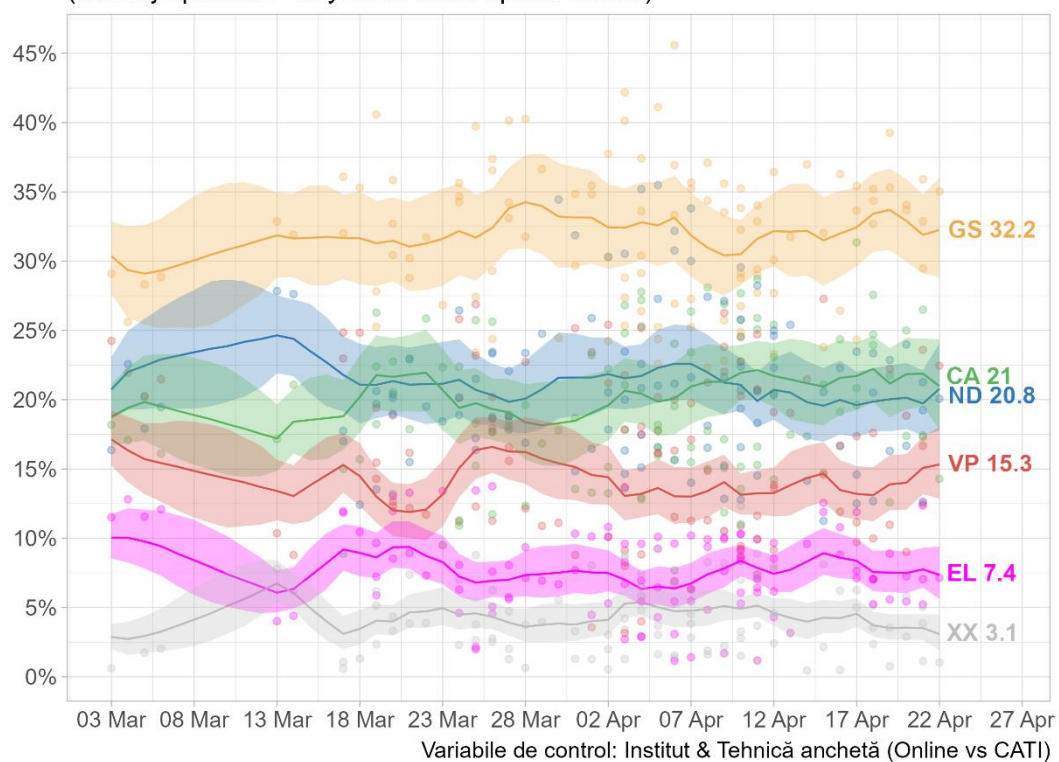
Prezidențiale 2025 (T1): Dinamica intențiilor de vot

(sondaje publice - modelare mlogit, # zile liniar + polinomial2&3)



Variable de control: Institut & Tehnică anchetă (Online vs CATI)

Prezidențiale 2025 (T1): Dinamica intențiilor de vot (sondaje publice - bayesian state-space model)



Au schimbat dezbaterile șansele percepute ale candidaților?

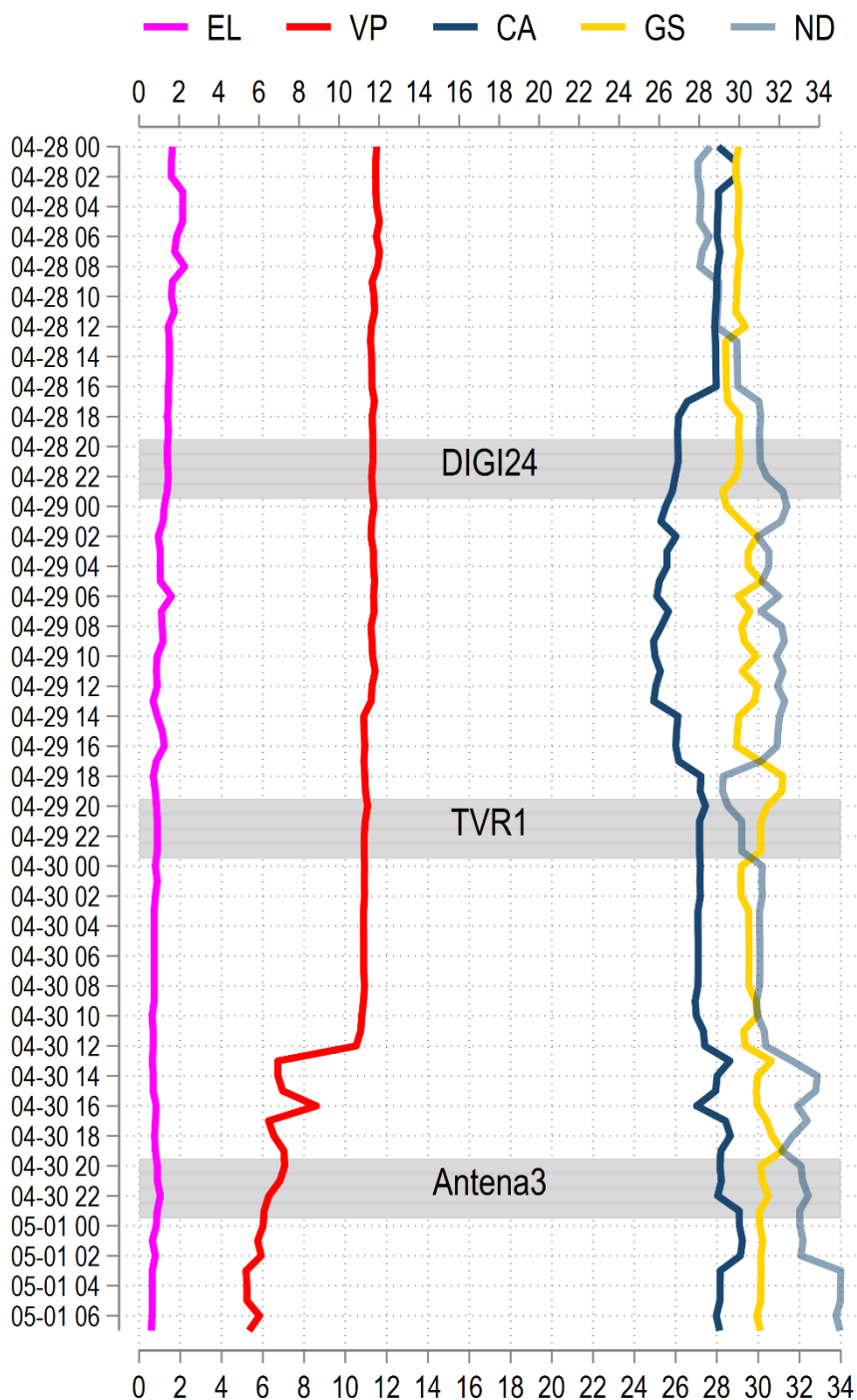
Data: 2025-05-01

Cuvinte cheie: alegerile-prezidentiale-2025, efectul-dezbaterilor

Pentru că date mai potrivite nu am, putem analiza efectul posibil al dezbaterilor uitându-ne la șansele percepute ale candidaților de a câștiga alegerile prezidențiale conform cotelor de la pariuri. Folosind datele furnizate de Polymarket, la interval de o oră, putem face o astfel de estimare. Atenție, cotele indică dinamica șanselor percepute de a câștiga alegerile din perspectiva unui public informat. La nivelul alegătorilor, respectiv votanților, lucrurile pot sta diferit. Câteva concluzii, repet, în termeni de șanse percepute de a câștiga alegerile (nu procente ce vor fi obținute la turul 1):

- Înainte de dezbateri, George Simion, Crin Antonescu și Nicușor Dan aveau șanse relativ apropiate.
- După dezbateri, șansele lui George Simion și Crin Antonescu au rămas aproximativ aceleași, iar șansele lui Nicușor Dan au crescut ușor.
- Șansele percepute ale celorlalți doi candidați, Victor Ponta și Elena Lasconi, s-au înjumătățit. În cazul lui Victor Ponta scăderea a avut loc preponderent înainte de dezbaterile de la Antena 3 (decizia de a nu participa?). În cazul Elenei Lasconi scăderea a fost relativ constantă.

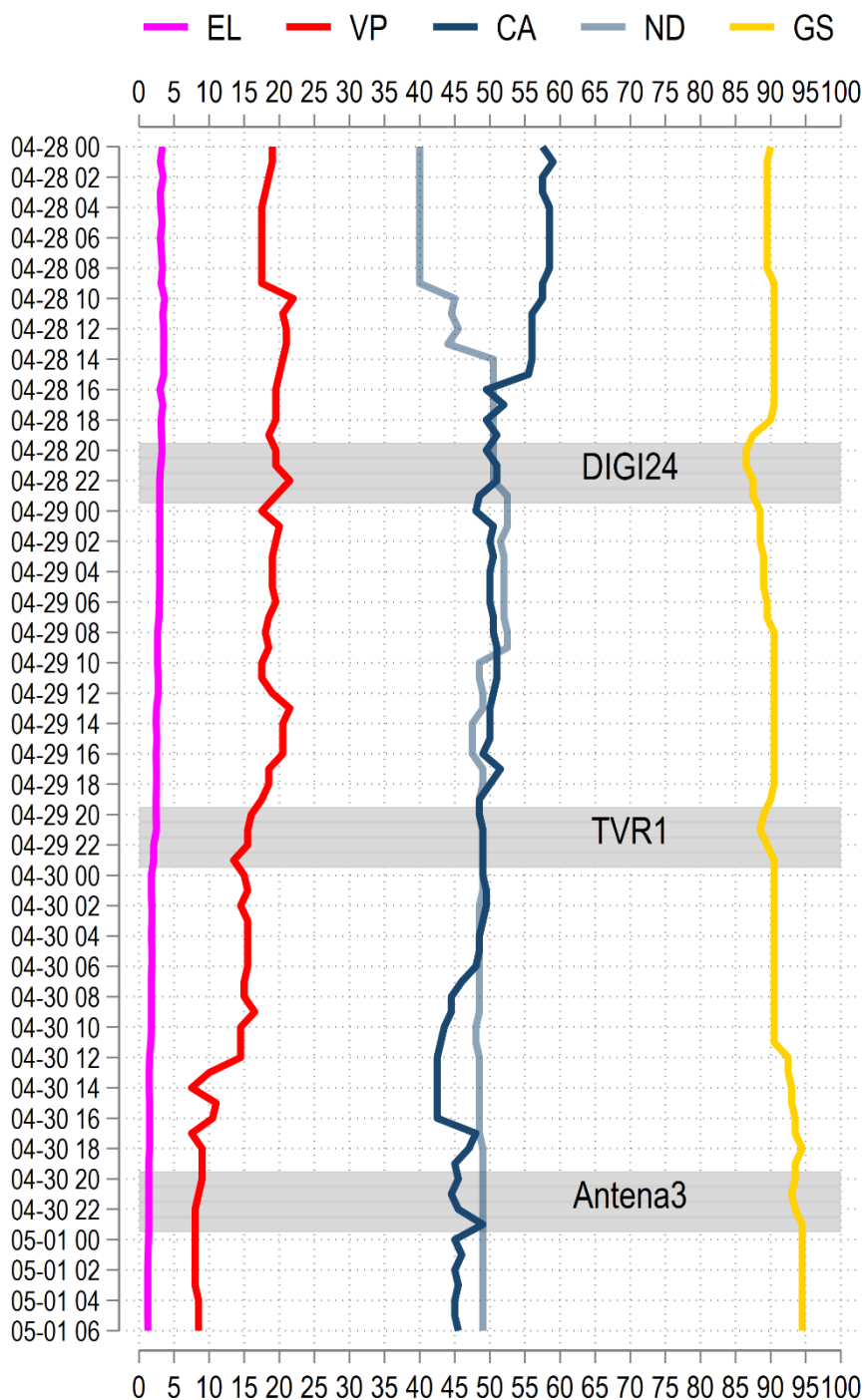
Dezbateri Prezidențiale 2025 - Cote pariuri - șansele percepute de a ajunge președinte -



Pun mai jos un grafic similar care indică dinamica șanselor candidaților de a ajunge în turul 2. Concluziile sunt următoarele:

- Șansele lui George Simion (90-95%) sunt pe un trend ușor ascendent care nu pare să fie influențat de dezbateri.
- Șansele lui Crin Antonescu (55-45%) sunt pe un trend ușor descendent care pare să fie influențat puțin de dezbateri.
- Șansele lui Nicușor Dan sunt relativ stabile în jurul a 50%, cu un impact minor al dezbaterilor.
- Șansele lui Victor Ponta (20-7%) s-au redus cu aproape două treimi, scăderea fiind ceva mai accentuată după prima dezbatere.
- Șansele Elenei Lasconi (4-2%) s-au redus la jumătate, panta fiind relativ constantă.

Dezbateri Prezidențiale 2025 - Cote pariuri - șansele percepute de a ajunge în T2 -



Alegerile prezidențiale 2025, turul 1: o predicție a intențiilor de vot

Data: 2025-05-02

Cuvinte cheie: alegerile-prezidentiale-2025, modelarea-intențiilor-de-vot, predicții

În postările anterioare despre dinamica intențiilor de vot am analizat strict datele din sondajele publice cu scopul de a simplifica informația și a prezenta tendințele indicate de acestea. În alte postări am analizat variația estimărilor în funcție de diferiți factori (firma care a realizat sondajul, eventualele conflicte de interese, respectiv modalitatea de colectare a datelor) cu scopul de a arăta că estimările produse de sondaje sunt, în unele cazuri, puternic dependente de acești factori. Folosesc aici rezultatele acestor analize, dar nu numai, pentru a produce o predicție a intențiilor de vot în cazul alegerilor de duminică.

Pentru realizarea predicției am combinat mai multe predicții, produse prin diferite metode, toate bazate pe datele din sondajele publice. Probabil ar fi fost utilă includerea și a altor tipuri de date (indicatori economici, cote pariuri, indicatori asociați mediului online). Predicția finală este o medie a patru tipuri de predicții, majoritatea construite pornind de la estimările din sondajele publice. Cele patru tipuri de predicții sunt următoarele:

- o medie a estimărilor realizate de firmele de sondare care au un grad relativ mai mare de încredere; am folosit doar sondajele cele mai recente;

- o modelare statistică a dinamicii intențiilor de vot pornind de la toate sondajele publice; aici am combinat predicțiile obținute cu ajutorul mai multor tipuri de analize: regresie logistică multinominală (liniar, non-liniar și combinații) și „bayesian state-space model”;
- folosind datele brute din trei seturi de date, două CATI (IRES și MMT) și unul online (MKOR), am construit și evaluat o serie de variante de ponderări, asociate unor scenarii diferite; am folosit media a două dintre predicțiile obținute; principala limită a acestui demers o constituie faptul că două dintre baze nu acoperă populația în vârstă (65+);
- folosind aceleași seturi de date brute, am realizat o modelare statistică de tip multi-nivel cu post-stratificare; desigur, ar fi fost de dorit ca toate bazele să acopere și populația vârstnică; o altă limită este numărul redus de cazuri; am folosit media mai multor predicții.

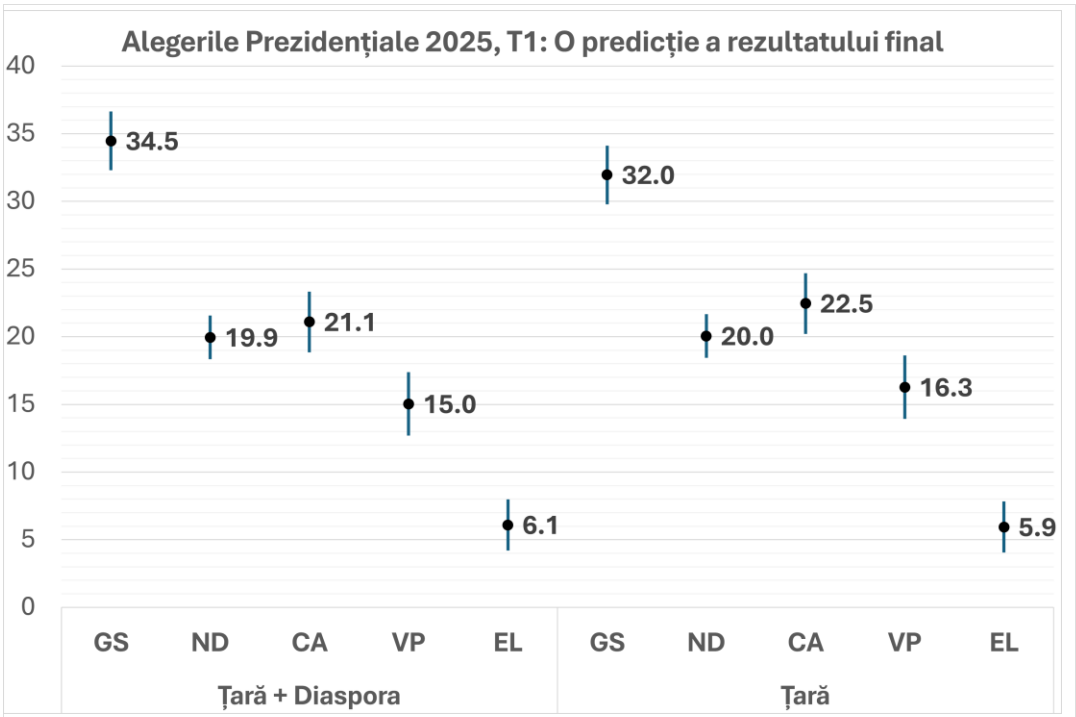
Majoritatea sondajelor publice acoperă doar populația din România. Pentru a produce o predicție care să includă și diaspora, am folosit distribuția voturilor de la turul 1 al alegerilor prezidențiale din 2024, estimări ale intențiilor de vot ale alegătorilor din diaspora conform datelor din sondajul AtlasIntel realizat în perioada 17-22 aprilie și estimările mele cu privire la împărțirea probabilă a voturilor.

Toate informațiile / estimările / aproximările cu care am lucrat au un anumit grad de incertitudine. Prin urmare, nu am folosit valori exacte ci distribuții asociate acestor valori. Am considerat că distribuțiile sunt de tip normal (formă de clopot). Prin faptul că folosește distribuții, modelarea ține cont de incertitudinile asociate valorilor de input. Desigur, distribuțiile au grade diferite de incertitudine. Le-am estimat folosind variațiile dintre estimările medii produse de cele patru metode menționate mai sus. Un model care nu ține cont de aceste aspecte riscă să producă estimări care (1) sunt prea sigure în relație cu informația disponibilă sau, dimpotrivă, (2) sunt prea imprecise, deci prea

puțin informative sau (3) asociază același grad de incertitudine a estimării tuturor candidaților.

Foarte probabil, modelul de agregare a informațiilor descris pe scurt mai sus nu va produce estimări perfecte. Faptul că informațiile folosite provin dominant din sondajele publice reprezintă o limită majoră (mai ales că nu am avut acces la bazele de date, cu mici excepții). Cu toate acestea, mă aștept ca scorurile finale să fie în general în interiorul intervalelor de încredere. Cu cât un scor va fi mai aproape de centrul intervalului de încredere, cu atât predicția va fi mai bună.

Pentru a nu deruta cititorii mai puțin obișnuiți cu grafice de tip distribuție de densitate, prezint rezultatul modelării într-o formă grafică simplă: cercurile pline indică predicțiile, iar liniile indică incertitudinea acestora, adică intervalul în care ne așteptăm, cu o probabilitate de 95%, să se situeze rezultatul final. În partea din dreapta a graficului apar predicțiile la nivel național (fără diaspora), iar în stânga predicțiile pentru național plus diaspora.



Pe scurt, cu referire doar la predicția la nivel național plus diaspora, mesajul graficului e următorul:

- pe primul loc va fi, aproape sigur, George Simion cu un scor de aproximativ 34-35% (intervalul de încredere e 32-37%);
- locurile doi și trei vor fi ocupate de Crin Antonescu și Nicușor Dan, probabil în această ordine, cu un scor de aproximativ 21% (Antonescu), respectiv 20% (Dan); rezultatele obținute nu permit să distingem între pozițiile celor doi candidați (intervalele de încredere se suprapun în mare parte); balanța poate fi înclinată într-o parte sau alta în funcție de distribuția deciziilor alegătorilor care s-au decis sau se vor decide zilele acestea, duminică inclusiv; cum, de obicei, numărul lor nu e deloc mic (~ 15-20%), oricare dintre cei doi candidați poate ajunge pe locul doi funcție de (1) mobilizarea și votul real al votanților PSD+PNL+UDMR, (2) o eventuală supra-mobilizare a alegătorilor în favoarea lui Nicușor Dan și (3) preferințele reale ale votanților din diaspora și, evident, numărul acestora;
- pe locul patru va fi mai probabil Victor Ponta cu un scor de aproximativ 15%; scorul final depinde mult de votul real al votanților PSD; dacă va fi preferat de o pondere mai mare a votanților PSD decât cea declarată în sondaje, scorului lui Ponta va fi ceva mai mare (și al lui Simion și/sau Antonescu ceva mai mic); estimarea asociată acestui candidat are un grad de incertitudine ridicat; incertitudinea este cauzată de diferența mare dintre estimările produse de sondajele telefonice comparativ cu sondajele online;
- pe locul cinci se va situa Elena Lasconi cu un scor de aproximativ 6%.

Prezidențiale 2025, T1: dinamica ratei de prezență

Data: 2025-05-04

Cuvinte cheie: alegerile-prezidentiale-2025, rata-de-prezentat-la-vot

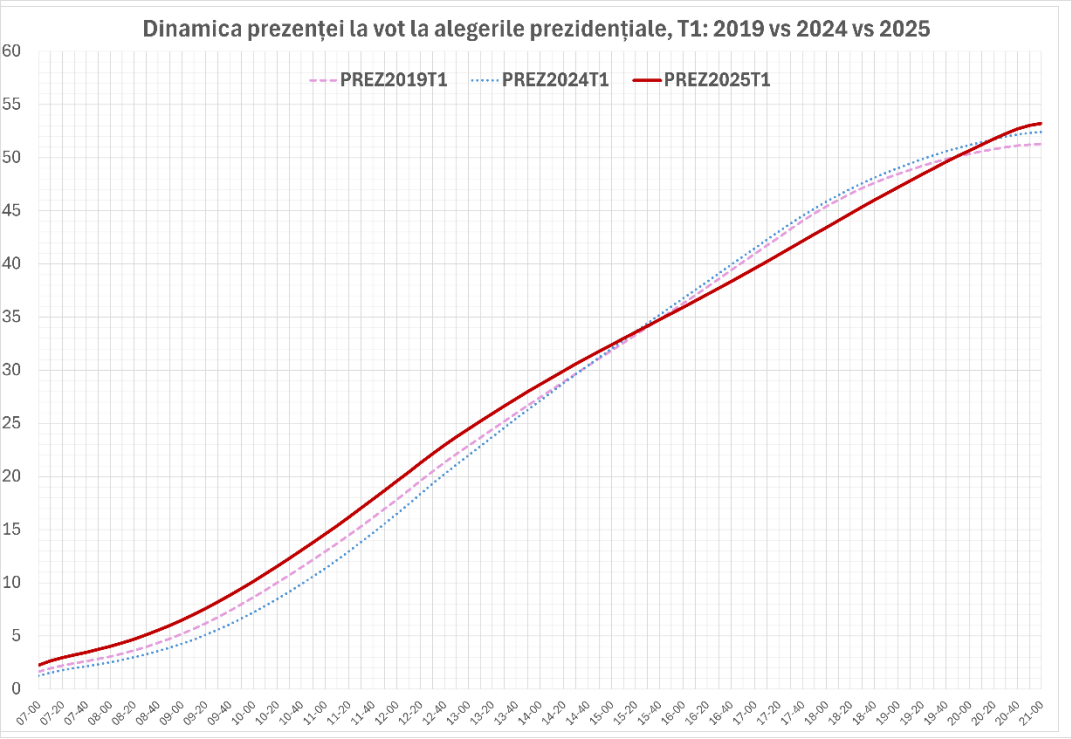
Un grafic cu rata de prezență la interval de 10 minute, turul 1 al alegerilor prezidențiale din 2025 vs. 2024 vs. 2019. Voi schimba graficul aici, din când în când.

11:00 Pe baza datelor de până la 11, prezența finală probabilă va fi 55-57%.

16:00 Ok, ne-am răcorit (curbele s-au apropiat), ne liniștim puțin, apoi vin forțe proaspete (valul de seară, destul de probabil)!

18:00 Cum am zis mai devreme. Să vedem dacă au mai rămas resurse pentru valul visat!

20:00 Ceva, ceva, dar insuficient ... pare că am fost nițel optimist.

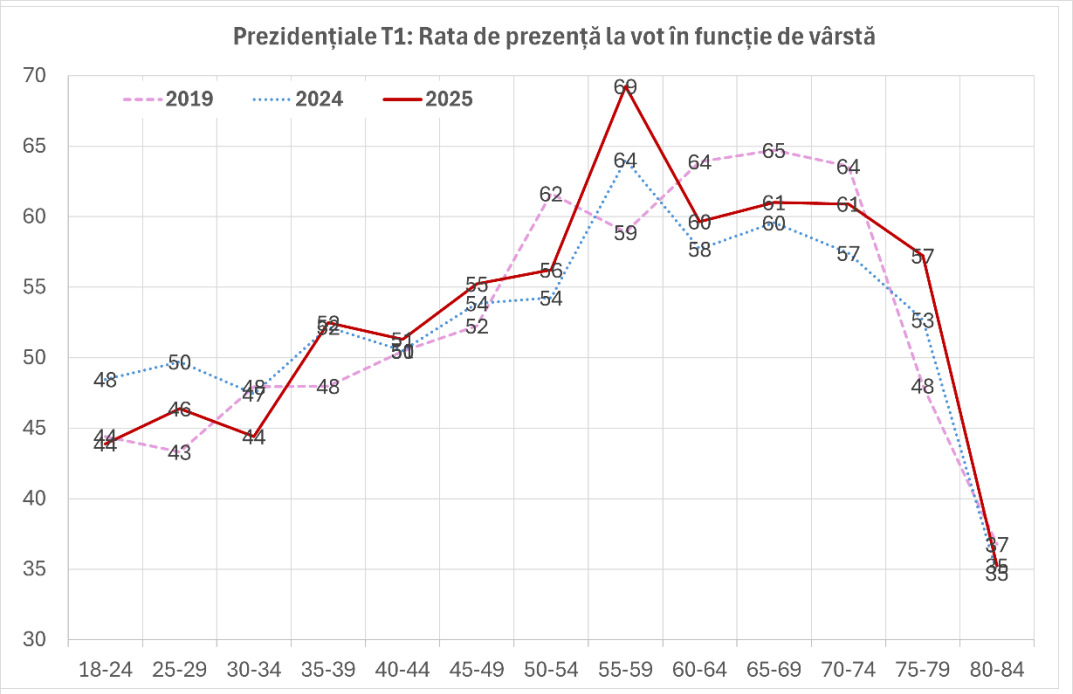


Prezidențiale 2025, T1: rata de prezență la vot în funcție de vârstă

Data: 2025-05-04

Cuvinte cheie: alegerile-prezidentiale-2025, rata-de-prezenta-la-vot

Pun aici o estimare a ratei de prezență la vot în funcție de vârstă. Pentru estimarea ratelor am folosit date AEP (numărul votanților), INS (populația rezidentă) și Eurostat (numărul emigranților). Desigur, ar fi fost mai simplu dacă AEP ar fi făcut publică structura pe vârste a alegătorilor, nu a doar a votanților. Iar estimările din grafic ar fi fost mai apropiate de realitate. Nu mai reiau toate explicațiile. Cine dorește, poate citi postarea anterioară (prezidențiale 2024) pe aceeași temă [aici](#).



Calitatea estimărilor sondajelor: Prezidențiale 2025, T1

Data: 2025-05-06

Cuvinte cheie: alegerile-prezidentiale-2025, performanta-sondajelor

Cu ocazia alegerilor din toamna anului trecut am realizat două analize cu privire la calitatea estimărilor produse de sondaje (vezi [aici](#) și [aici](#)). Astfel de analize sunt importante din multe puncte de vedere și pentru diferite tipuri de actori (media, publicul larg, institutele de sondaje, beneficiarii actuali și potențiali ai acestora). Dat fiind faptul că realitatea măsurată de sondaje este adesea fluidă, e important să distingem între sondajele realizate în săptămâna anterioară alegerilor și cele realizate în ziua alegerilor. Suplimentar, voi analiza și performanța unor modele de predicție bazate pe sondaje. Pentru a ține cont de aceste aspecte, am separat analiza în trei secțiuni: (1) sondaje realizate cu cel mult 10 zile înainte de data alegerilor, (2) modelarea estimărilor / datelor din sondaje și (3) sondaje realizate în ziua votului. Înainte de a trece la prezentarea rezultatelor analizei, sunt necesare câteva precizări metodologice, unele ceva mai tehnice.

Am considerat doar candidații care, anterior alegerilor, în sondaje, aveau un scor de cel puțin 5%: George Simion, Crin Antonescu, Nicușor Dan, Victor Ponta și Elena Lasconi (restul candidaților au format categoria „alt candidat”). Majoritatea sondajelor au fost realizate la nivel național (au acoperit doar populația din țară), prin urmare am comparat estimările produse de sondaje cu rezultatul final la nivel de țară. Situațiile care au la

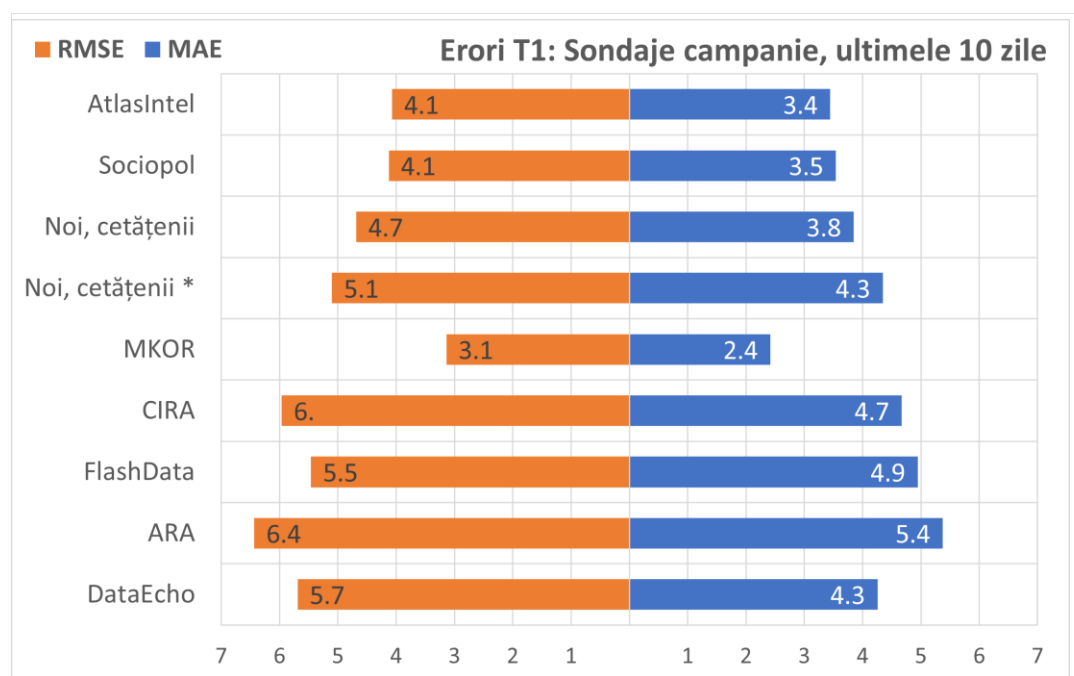
final semnul * indică faptul că sondajul / predicția a fost la nivel țară + diaspora, comparația fiind făcută cu rezultatul oficial de la același nivel.

Pentru a măsura calitatea estimărilor am folosit două măsuri: MAE și RMSE. MAE (Mean Average Error) se referă la media erorilor absolute, eroarea fiind definită ca diferența absolută dintre procentul estimat de sondaj și procentul obținut conform BEC. O astfel de măsură indică distanța medie absolută dintre estimările produse de sondaje și rezultatele finale. MAE tratează toate erorile la fel, indiferent de direcția lor și este mai puțin influențată de existența unor valori extreme. Simplu spus, distinge mai greu între următoarele două situații: (A) o firmă de sondaje care face o eroare mare de predicție în cazul unui candidat și multe erori mici în cazul celorlalți și (B) o firmă de sondaje care face erori medii în cazul tuturor candidaților. RMSE (Root Mean Squared Error), rădăcina medie pătratică a erorilor, se calculează ca radical de ordinul doi din suma erorilor ridicate la pătrat. Comparativ cu MAE, prin faptul că penalizează mai mult erorile mari, RMSE distinge ceva mai bine între cele două tipuri de situații. Prin urmare, dacă RMSE este mai mare decât MAE indică faptul că unul dintre candidați a fost estimat de semnificativ mai incorect. În cazul ambelor măsuri, o valoare mai mare indică o diferență (~ eroare) medie mai mare între estimări și rezultatul oficial. **Diferențele mai mari de trei puncte procentuale pot fi considerate problematice** (două puncte procentuale în cazul sondajelor realizate pe eșantioane mari - minimum 2000 „cazuri efective” = numărul cazurilor recalculat pentru situația în care eșantionul a fost de tip SRS (Simple Random Sample), adică simplu aleator.

Estimările sondajelor realizate în ultimele 10 zile

Aproape toate sondajele realizate cu cel mult 10 zile înainte de data alegerilor au avut erori medii mai mari de pragul de trei puncte procentuale (pp). Într-un singur sondaj - [MKOR](#) - erorile au fost apropiate de prag: MAE = 2.4 (puțin sub prag) și RMSE = 3.1 (puțin peste prag). Pentru transparență, precizez următoarele: (1)

MKOR este o companie relativ nouă, axată pe cercetarea de marketing; (2) populația măsurată în sondaj a fost 18-64 ani, preponderent din urban (publicul mai degrabă comercial), deci nu a acoperit toată populația cu drept de vot; (3) sondajul a fost de tip online; (4) la cererea lor, le-am oferit consultanță neplătită, strict pentru sondajele politice. Sondajele AtlasIntel și Sociopol au produs estimări destul de bune, totuși peste pragul de trei pp. Erorile medii cele mai mari (4-6 pp) s-au înregistrat în cazul sondajelor de tip CATI realizate de ARA și CIRA, respectiv a celor online realizate de DataEcho, FlashData și „Noi, cetățenii” (estimarea cu * e la nivel țară + diaspora, deci comparația e cu rezultatul la acel nivel). Fiind vorba de sondaje realizate prin tehnici diferite, e mai puțin probabil ca erorile apărute să aibă drept principală cauză tehnica de sondaj (telefonice sau online).



Valorile prezentate în graficul anterior sunt medii, deci pot să ascundă diferențele dintre estimările punctuale. Pentru o imagine completă ar trebui să ne uităm, la nivelul fiecărui candidat, și la diferența dintre estimarea minimă și cea maximă. Scorurile minime și maxime obținute de sondajele din ultimele 10 zile arată așa:

- George Simion (scor final în țară 38.7%): de la 27% (ARA) până la 34% Sociopol, adică o amplitudine de 7 pp.
- Nicușor Dan (scor final în țară 20.5%): de la 19% (Sociopol) până la 27% (DataEcho), adică o amplitudine de 8 pp.
- Crin Antonescu (scor final în țară 21.6%): de la 20% (CIRA) până la 27% (FlashData), adică o amplitudine de 7 pp.
- Victor Ponta (scor final în țară 14.3%): de la 8% (FlashData) până la 22% (Sociopol), adică o amplitudine de 14 pp.
- Elena Lasconi (scor final în țară 2.6%): de la 5% (Sociopol) până la 10% (ARA), adică o amplitudine de 5 pp.

Predicțiile pe baza estimărilor din sondajele publice

Ce se întâmplă dacă facem o medie a estimărilor din sondaje, dacă modelăm datele din sondaje, respectiv dacă realizăm un model care ia în calcul, pe lângă estimările din sondaje, și alte informații? Care dintre aceste variante produce predicții mai bune? Pentru a răspunde la această întrebare am procedat astfel:

- Am calculat media simplă a estimărilor produse de sondajele realizate cu cel mult 10 zile înainte de alegeri.
- Am calculat aceeași medie în cazul sondajelor CATI și a celor online.
- Am inclus în comparație estimările finale produse de [EuropeElects](#) și [Politico](#) pe baza datelor din sondajele publicate.

- Am inclus predicția realizată de mine și publicată pe blog în ziua de vineri (vezi [aici](#)). Am inclus ambele predicții (cu și fără voturile din diaspora).

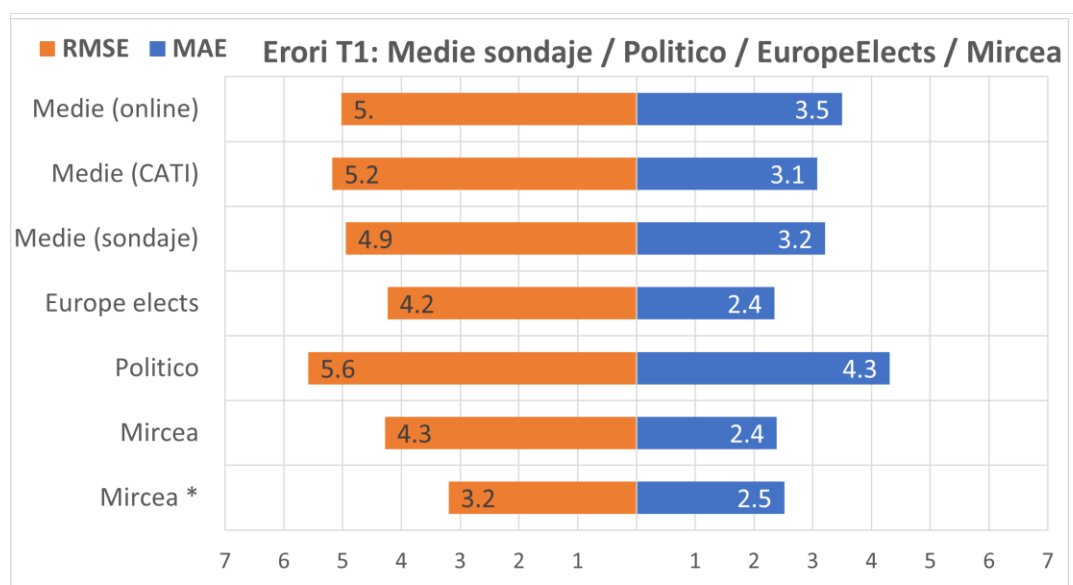
Medii și modelări ale estimărilor din sondaje

O medie simplă a sondajelor din ultimele 10 zile (Medie sondaje) a produs o estimare care depășește pragul de trei pp. Media erorilor diferă puțin în funcție de tehnica de sondaj cu o mențiune importantă: pe baza măsurii MAE sondajele CATI stau puțin mai bine, iar pe baza măsurii RMSE eroarea e mai mică în cazul sondajelor online. Prin urmare, cel puțin în acest caz, e de preferat să combinăm estimările produse de cele două tipuri de sondaje. Însă, chiar și așa, erorile sunt mari. Sondajele MKOR (online), Sociopol (CATI) și AtlasIntel (online) au produs estimări cu o eroare mai mică comparativ cu media sondajelor. În concluzie, o estimare bazată pe media tuturor sondajelor nu e întotdeauna mai bună decât estimările individuale ale unora dintre sondaje. Important e să identificăm corect sondajele de încredere, ceea ce de altfel am și făcut atunci când am folosit pentru predicție doar sondajele cu un grad mai mare de încredere.

Estimările produse prin agregarea sondajelor, Politico și EuropeElects, au avut erori foarte diferite. Deși a depășit pragul de trei pp, predicția realizată de EuropeElects a fost printre cele mai bune. De cealaltă parte, predicția realizată de Politico a avut cele mai mari erori.

Predicțiile realizate de mine, prin agregarea și modelarea estimărilor din sondaje și a trei baze brute, au avut cele mai mici erori (alături de predicția EuropeElects). Predicția care a vizat rezultatul final (Mircea *), adică țară plus diaspora, a avut cele mai mici erori dintre toate estimările bazate pe date agregate: MAE = 2.5 (sub prag) și RMSE = 3.2 (puțin peste prag). Această predicție are o calitate similară cu cea mai bună predicție din sondajele pre-electorale (sondajul MKOR). Simultan predicția (Mircea *) a fost mai bună decât predicțiile unora dintre sondajele realizate în ziua alegerilor (exit-poll-ul Avangarde și sondajul

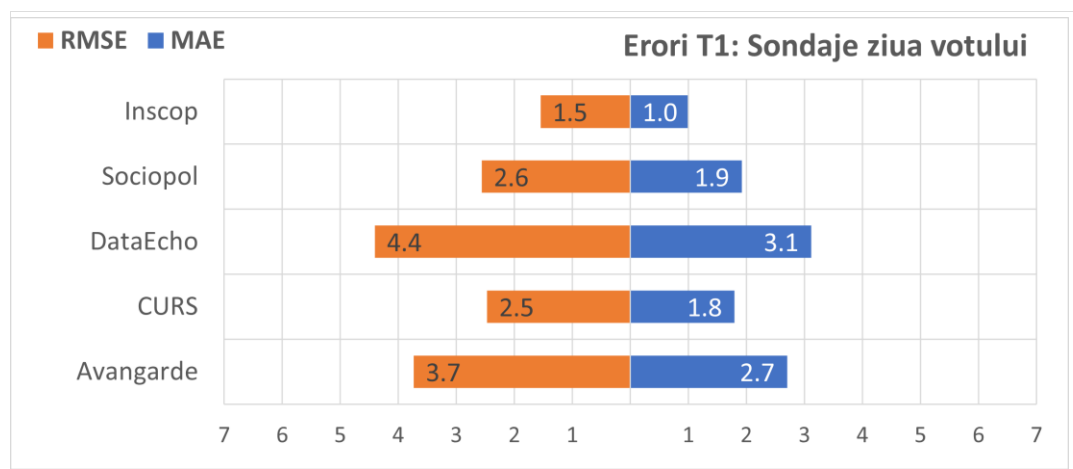
online DataEcho). Aceeași predicție a avut erori puțin mai mari comparativ cu predicțiile CURS (exit-poll) și Sociopol (sondaj telefonic în ziua votului), respectiv erori mai mari comparativ cu predicția Inscop (sondaj telefonic realizat în ziua votului). Predicția mea la nivel țară (fără diaspora) a avut erori ceva mai mari, similare cu predicția EuropeElects. Totuși, ambele predicții (Mircea și EuropeElects) au fost în medie mai bune decât predicțiile produse de aproape toate sondajele sau de media acestora. Prin urmare, dacă suntem înaintea de alegeri și dorim să știm care vor fi rezultatele alegerilor, ar trebui să ne uităm mai puțin la sondaje punctuale și mai mult la predicțiile realizate pe baza modelărilor estimărilor din sondaje.



Estimările sondajelor din ziua votului

Sondajele din ziua votului au produs, evident, cele mai bune predicții, deși, și în acest caz au existat firme de sondare care au avut erori mai mari decât ne-am fi așteptat. Cele mai mari erori medii au fost în cazul firmei DataEcho (online), urmată de Avangarde (exit-poll). Dintre toate sondajele, sondajul telefonic realizat de Inscop a avut, de departe, cea mai mică eroare medie.

Sociopol (telefonic) și CURS (exit-poll) s-au situat undeva între cele două extreme.



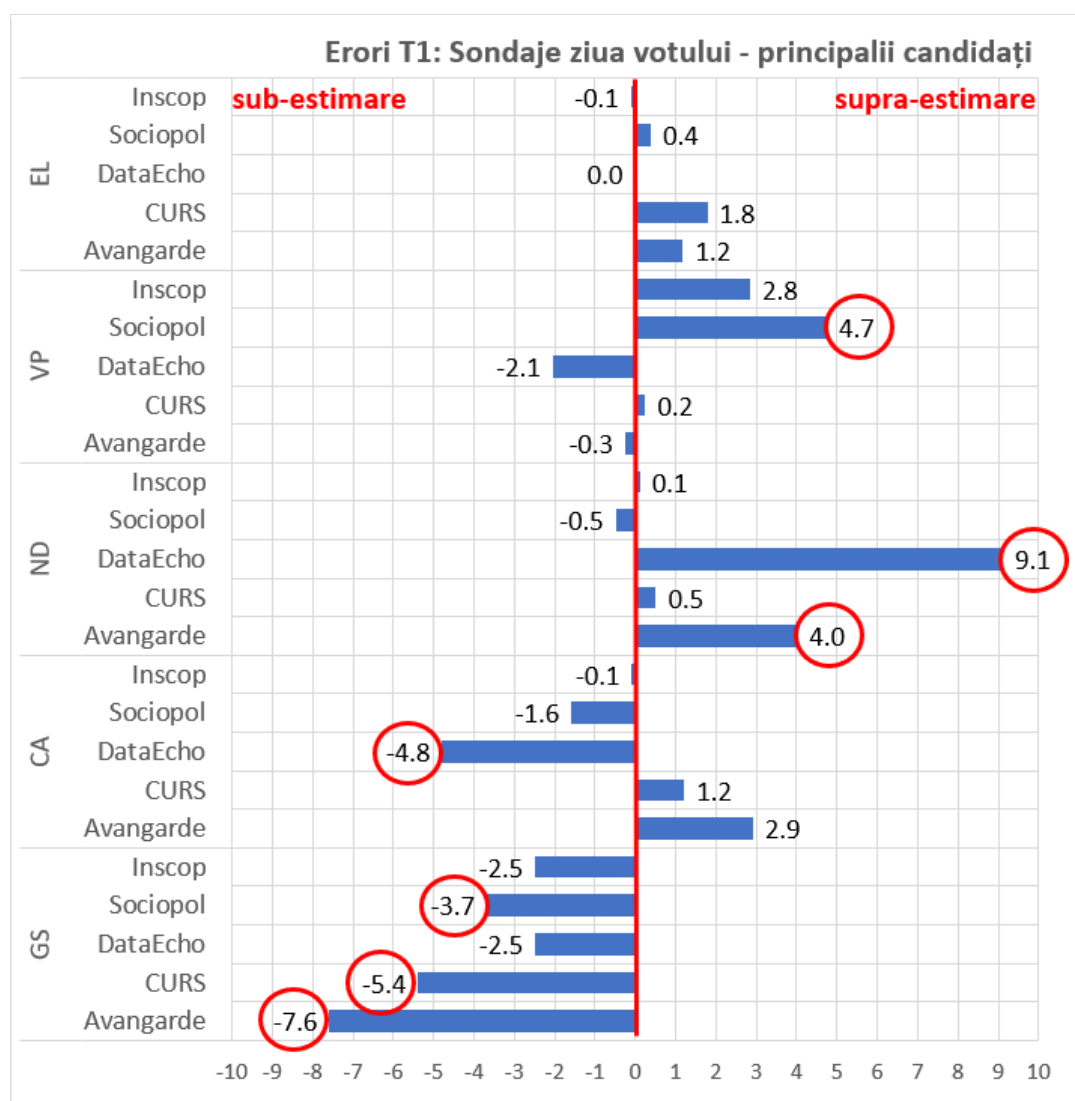
Sub/Supra-estimări în ziua votului

În graficele anterioare, distanța dintre estimări și realitate este calculată ca medie a unor diferențe absolute. Pentru o imagine completă, e necesar să ne uităm și la diferențele punctuale observate în cazul candidaților majori (cei care au obținut mai multe procente). Procedând astfel, putem vedea și sensul diferențelor, respectiv putem face unele inferențe cu privire la sursa erorilor. Astfel, putem observa dacă un candidat a fost sub-estimat sau supra-estimat în sondaje, respectiv dacă un eventual efect de acest tip variază sau nu în funcție de compania care a realizat sondajul și/sau tehnica de sondaj folosită. Prin sub-estimare mă refer la faptul că un candidat a obținut în sondaje o estimare a intenției de vot mai mică față de procentul înregistrat la alegeri (supra-estimarea se referă la fenomenul invers: un scor mai mare în sondaje comparativ cu scorul obținut la alegeri). Pentru a nu lungi prea mult analiza, prezint doar **evaluarea cu privire la sondajele realizate în ziua votului**. Din perspectiva candidaților, situația se prezintă astfel:

- **George Simion a fost sub-estimat de toate firmele de sondare.** Sub-estimarea minimă a fost de 2.5 puncte procentuale: Inscop (telefonice) și DataEcho (online). **Sub-estimarea maximă a fost de 7.6 puncte: Avangarde (exit-poll).** O sub-estimare similară s-a întâmplat anterior și în cazul lui Călin Georgescu. Foarte probabil, sub-estimarea lui George Simion are legătură cu efectul de dezirabilitate socială – o poziționare critică a mass-media la adresa lui George Simion a făcut ca o parte a votanților să-și ascundă intenția de vot. În același sens a acționat și faptul că PSD a cerut propriilor susținători să voteze cu un candidat care nu provine din rândurile propriului partid: o parte a votanților PSD au răspuns la întrebarea de vot din sondaje „Crin Antonescu” în timp ce în cabină au pus ștampila pe „Victor Ponta” sau pe „George Simion.”
- **Crin Antonescu a fost sub-estimat de DataEcho (online) și, la limita pragului, supra-estimat de Avangarde (exit-poll).** Dat fiind faptul unul dintre sondaje a fost online, iar celălalt la urne, cauza probabilă a sub-estimării, respectiv supra-estimării are legătură (și) cu distorsiunea introdusă de tehnica de sondaj folosită.
- **Nicușor Dan a fost mult supra-estimat (+8) de DataEcho (sondaj online) și supra-estimat de Avangarde (exit-poll).** Parțial, supra-estimarea de către Avangarde a candidaților Antonescu și Dan este o consecință a sub-estimării mari observate în cazul lui Simion.
- **Victor Ponta a fost supra-estimat de Sociopol cu aproape cinci puncte procentuale.** Supra-estimarea nu a fost la fel de mare precum cea din sondajele realizate în perioada de campanie. În sondajele realizate de Sociopol în campanie, Victor Ponta apărea sistematic cu o intenție de vot de 22-23%. La alegeri, Ponta a obținut 14.3%, deci cu 8 pp mai puțin.

Din perspectiva firmelor de sondare, situația arată așa:

- Avangarde l-a sub-estimat masiv pe George Simion (-7.6) și l-a supra-estimat pe Crin Antonescu (+2.9).
- DataEcho l-a supra-estimat masiv pe Nicușor Dan (+9.1) și l-a sub-estimat pe Crin Antonescu (-4.8).
- CURS l-a sub-estimat mult pe George Simion (-5.4).
- Sociopol l-a sub-estimat pe George Simion (-3.7) și l-a supra-estimat mult pe Victor Ponta (+4.7).
- În cazul Inscop, toate erorile au fost sub pragul de trei puncte procentuale. Au existat totuși unele abateri: sub-estimare în cazul lui George Simion (-2.5) și supra-estimare în cazul lui Victor Ponta (+2.8).



Câteva concluzii

Din nou, similar cu alegerile de anul trecut, erorile asociate candidatului „suveranist” au fost mai mari în cazul sondajelor realizate la ieșirea de la urne (ceea ce numim exit-poll). E puțin probabil ca eroarea asociată exit-poll-urilor să fie cauzată de eroarea de eșantionare sau de formularea întrebării de vot. O cauză mai probabilă este ponderea mare a refuzurilor în cazul

exit-poll-urilor. Din informațiile postate online (<https://www.rezultatevot.ro/embed/articol/44>), pare că ponderea refuzurilor la întrebarea de vot a fost de aproximativ 50%. Pe de altă parte, și în sondajul telefonic realizat de Inscop, cel cu predicțiile cele mai bune, numărul refuzurilor a fost aproape la fel de mare (45%). Prin urmare, ceea ce probabil a contat în primul rând nu a fost ponderea mare a refuzurilor ci mai ales modul în care au fost tratate refuzurile. Aproape sigur, una dintre cauzele care au dus la diferențele observate a fost modul în care statisticienii acestor firme de sondare au tratat refuzurile. Desigur, erorile au fost influențate și de tehnica de sondaj folosită.

Pe baza celor descrise mai sus, putem deduce că a existat o cauză generală care a condus la sub-estimarea candidatului George Simion. Foarte probabil e vorba despre ascunderea opțiunii de vot pe diferite căi: (1) rata diferită de acceptare a participării la sondaj, (2) rata diferită de declarare a opțiunii de vot și (3) unii votanți au declarat altă opțiune decât cea reală. Aproape sigur, un impact semnificativ au avut și modul în care au fost tratate refuzurile, respectiv tipul de sondaj (la urne vs. telefonic vs. online).

Despre transparența sondajelor

Un comentariu final, aparent marginal, dar extrem de important, cu privire la **transparența sondajelor din România**. Transparența este un criteriu important în funcție de care putem evalua calitatea unui sondaj și profesionalismul unui institut de sondaje. Prin transparență mă refer la tipul de informații care sunt făcute publice, alături de selecția de grafice, și la gradul de detaliere al acestor informații. În cazul sondajelor politice din România, cel mai adesea, sunt prezentate doar informații minimale, fără prea multe detalii și fără să fie acoperite o serie de informații care sunt foarte importante pentru cineva care dorește să analizeze calitatea datelor produse. Puține institute publică un raport pe propriul site, deci analizele pot fi făcute doar pe baza informațiilor care apar în media. În cazul sondajelor din ziua votului, lucrurile stau chiar și mai rău, cu excepția **Inscop** și **CURS** care au prezentat pe site-ul propriu câteva informații, fie ele și minimale. Nu am

găsit niciun alt raport pe paginile web ale celorlalte firme de sondare, prin urmare am vânat estimările apărute în media. O astfel de practică e problematică, cel puțin pentru faptul că, uneori, datele prezentate diferă de la o sursă media la alta. Problema nu e deloc una nouă. A fost semnalată în urmă cu zece ani de Bogdan Voicu într-o serie de [postări](#). S-a schimbat ceva între timp? Sigur da. Din păcate schimbarea a fost în rău. Dacă în urmă cu zece ani mai apărea pe ici pe colo o descriere sumară a metodologiei sondajelor din ziua votului, în prezent, uneori, nu mai avem nici măcar atât.

Prezidențiale 2025, T2: dinamica T2 anterior T1

Data: 2025-05-07

Cuvinte cheie: alegerile-prezidentiale-2025, dinamica-intențiilor-de-vot

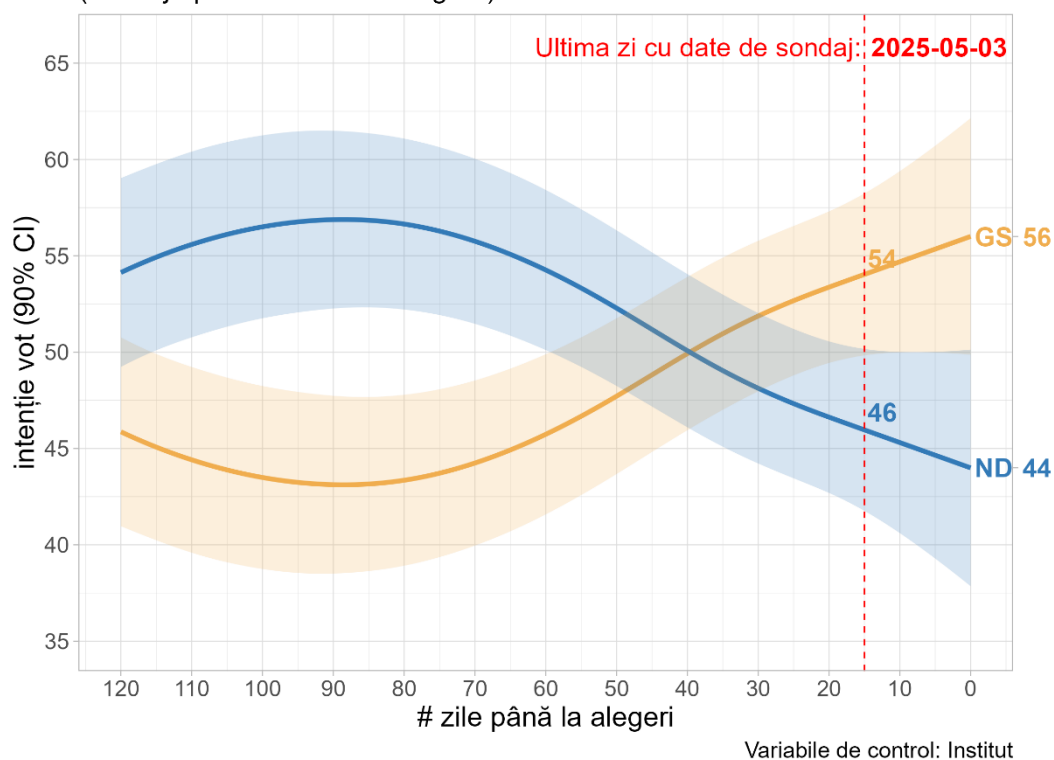
Fac aici o scurtă analiză în care prezint dinamica intențiilor de vot pentru turul 2 pornind de la estimările din sondajele „publice” realizate anterior turului 1. Pun publice între ghilimele pentru că nu toate sondajele folosite aici au și fost publicate în forma obișnuită. Astfel, am inclus în analiză și [estimările prezentate sub forma unui grafic](#) de Mirel Palada (Sociopol) pe contul său de FB (doar o parte a acestora au fost publicate anterior în varianta clasică). Am extras din grafic estimările și le-am adăugat la estimările preluate din sondajele publice. Am modelat tot setul de date folosind două tipuri de analiză statistică: GAM și bayesian state-space model. Prezint rezultatul modelărilor sub forma a două grafice. Acestea indică **dinamica probabilă a intențiilor de vot la nivel de țară (fără diaspora)**.

E important să menționez că, relativ la prima jumătate a perioadei analizate, cele două serii de estimări spun o poveste diferită, prin urmare, dinamica prezentată aici diferă de dinamica din graficul postat de Mirel Palada - însă, dinamica din ultima lună este similară. Aceasta se întâmplă și pentru că majoritatea estimărilor „oficiale,” altele decât cele făcute de Sociopol, au fost produse folosind sondaje online. Am văzut într-o [analiză anterioară](#) că sondajele online tind să-l supra-estimeze pe ND. Majoritatea sondajelor au prezentat trei estimări: intenția de vot pentru GS, ND și o categorie care însumează indecizii, non-răspunsurile,

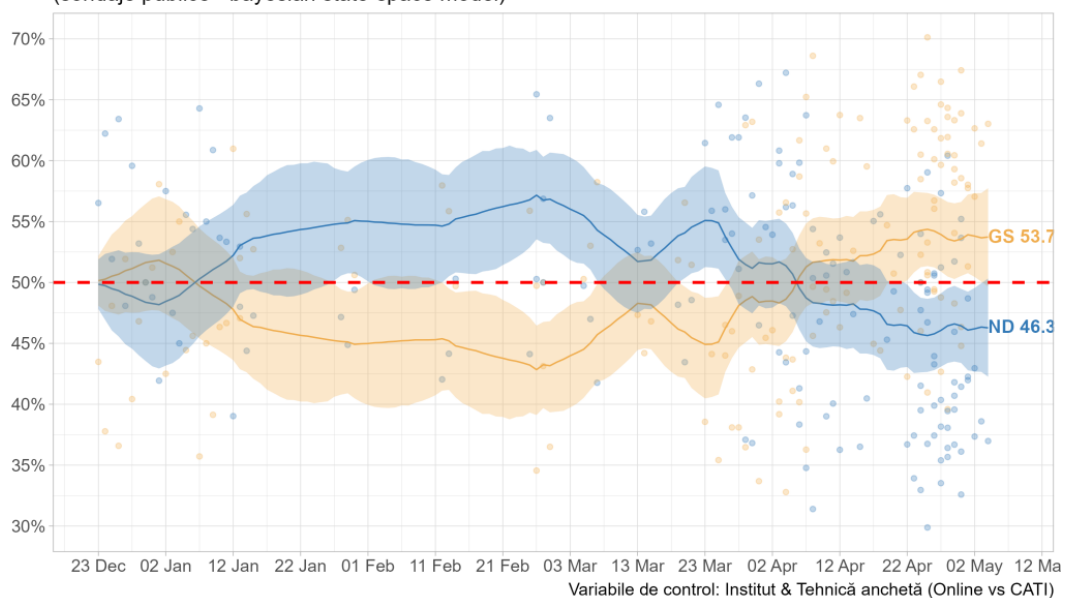
uneori și non-votanții. Pentru ca estimările să fie comparabile, le-am recalculat ca **procente din răspunsuri valide** (cei care au indicat un candidat).

Concluzii: Înaintea primului tur, distanța dintre cei doi finaliști era de aproximativ 8 puncte procentuale (pp). Luând în calcul și distorsiunea sondajelor în defavoarea lui George Simion (vezi efectul de dezirabilitate socială identificat într-o [analiză anterioară](#)), estimez că distanța era în realitate mai mare de 8 pp, probabil în jur de 12-14 pp. Dacă tendința observată relativ la ultima lună continuă și în următoarele două săptămâni, e mai probabil ca distanța dintre candidați să crească. Ambele modele statistice indică acest lucru, cu unele diferențe: la modelul GAM panta de creștere a lui George Simion e clară; la modelul bayesian, în ultimele două săptămâni, panta s-a aplatizat (în cazul estimărilor Sociopol aplatizarea a început în urmă cu o lună). Desigur, dacă apar evenimente majore și/sau dacă are loc o mobilizare masivă a alegătorilor, ambele dominant în favoarea unuia dintre candidați, distanța se va modifica. Și nu în ultimul rând, contează ce se va întâmpla și în diaspora.

Prezidențiale 2025 (T2): Dinamica intențiilor de vot (sondaje publice - modelare gam)



Prezidențiale 2025 (T2): Dinamica intențiilor de vot (sondaje publice - bayesian state-space model)



Prezidențiale 2025: dinamica T2

Data: 2025-05-15

Cuvinte cheie: alegerile-prezidentiale-2025, dinamica-intențiilor-de-vot

Fac aici o scurtă analiză în care prezint dinamica intențiilor de vot pentru turul 2 pornind de la estimările din sondajele „publice” realizate până la acest moment. Pun publice între ghilimele pentru că nu toate sondajele folosite aici au și fost publicate în forma obișnuită. Astfel, am inclus în analiză și [estimările prezentate sun forma unui grafic](#) de Mirel Palada (Sociopol) pe contul său de FB (doar o parte a acestora au fost publicate anterior în varianta clasică). Am extras din grafic estimările și le-am adăugat la estimările preluate din sondajele publice. De asemenea, am inclus și ultimele estimări apărute în (social) media, fără să fie prezentat un grafic sau un tabel cu rezultatele (Inscop, Sociopol). Am modelat setul de date folosind aceleași tipuri de analiză statistică. Pentru modelul bayesian am estimat două variante, cu și fără variabila de control modalitatea de colectare a datelor (CATI vs. Online).

E important să menționez că, relativ la prima jumătate a perioadei analizate, cele două serii de estimări spun o poveste diferită, prin urmare, dinamica prezentată aici diferă de dinamica din graficul postat de Mirel Palada - însă, dinamica din ultima lună este similară. Aceasta se întâmplă și pentru că majoritatea estimărilor „oficiale,” altele decât cele făcute de Sociopol, au fost produse folosind sondaje online. Am văzut într-o [analiză anterioară](#) că sondajele online tind să-l supra-estimeze pe ND. Majoritatea

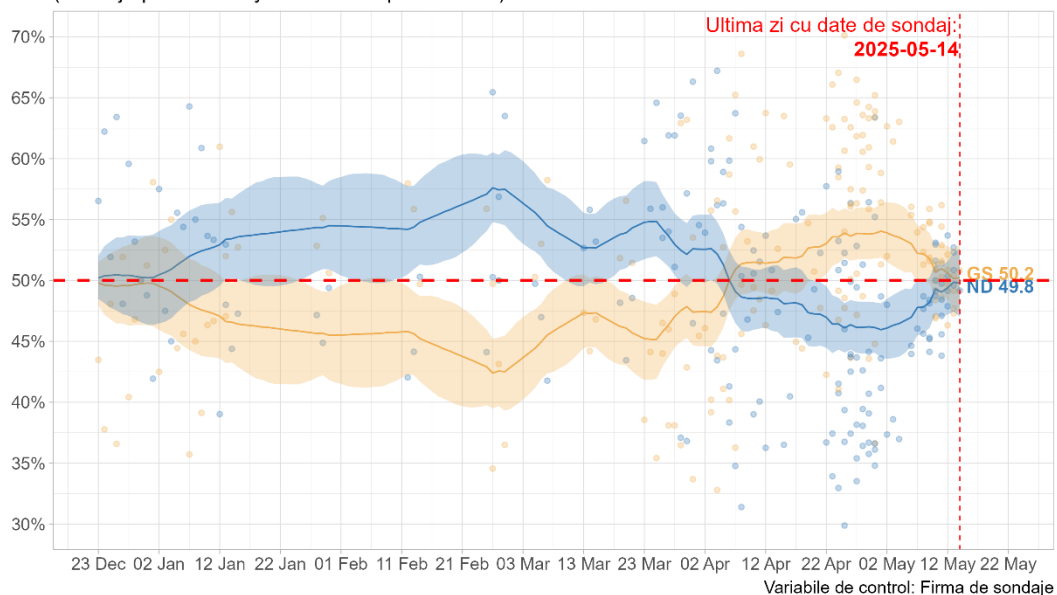
sondajelor au prezentat trei estimări: intenția de vot pentru GS, ND și o categorie care însumează indecisiile, non-răspunsurile, uneori și non-votanții. Pentru ca estimările să fie comparabile, le-am recalculat ca **procente din răspunsuri valide** (cei care au indicat un candidat).

Graficele de mai jos indică **dinamica probabilă a intențiilor de vot la nivel de țară (fără diaspora)**. Estimările din grafice nu reprezintă (neapărat) o predicție a rezultatului alegerilor din țară ci o sinteză a estimărilor din sondajele publice. Voi publica mâine o predicție bazată pe estimările agregate ale sondajelor plus alte informații și analize statistice.

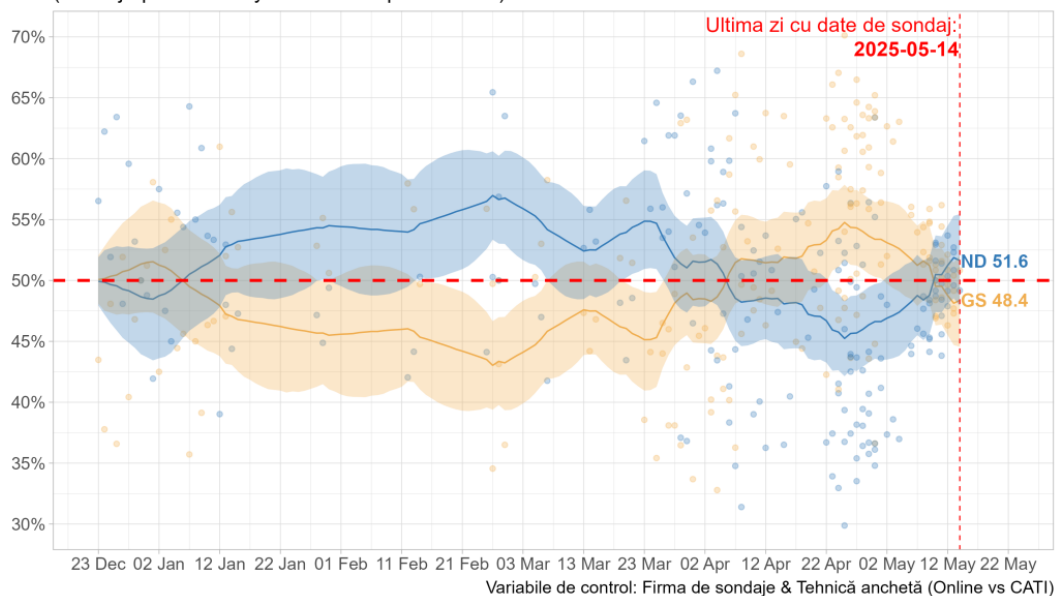
Reamintesc faptul că în proximitatea primului tur, distanța dintre cei doi finaliști era de aproximativ 8 puncte procentuale (pp). Luând în calcul și distorsiunea sondajelor în defavoarea lui George Simion (vezi efectul de dezirabilitate socială identificat într-o [analiză anterioară](#)), distanța era în realitate mai mare de 8 pp, probabil în jur de 12-14 pp.

Chiar dacă estimările produse de modele par să fie diferite, din punct de vedere statistic sunt similare. La data ultimului sondaj (14 mai), distanța dintre cei doi candidați era în marja de eroare, deci oricare dintre aceștia putea fi pe primul loc în acel moment. Funcție de cum anume mai exact au fost produse estimările din sondajele publicate - dacă firmele de sondare au ținut sau nu cont de efectul de (in)dezirabilitate socială, dacă și cum anume au ponderat datele brute, cum anume au tratat indecisiile - lucrurile pot sta diferit cu până la două-trei puncte procentuale (într-un sens sau altul). Dat fiind faptul că tendința din ultimele două săptămâni este în favoarea lui Nicușor Dan, este mai probabil ca acesta să câștige alegerile din țară. Întrebarea e dacă avansul obținut în țară va fi suficient de mare astfel încât să câștige și alegerile.

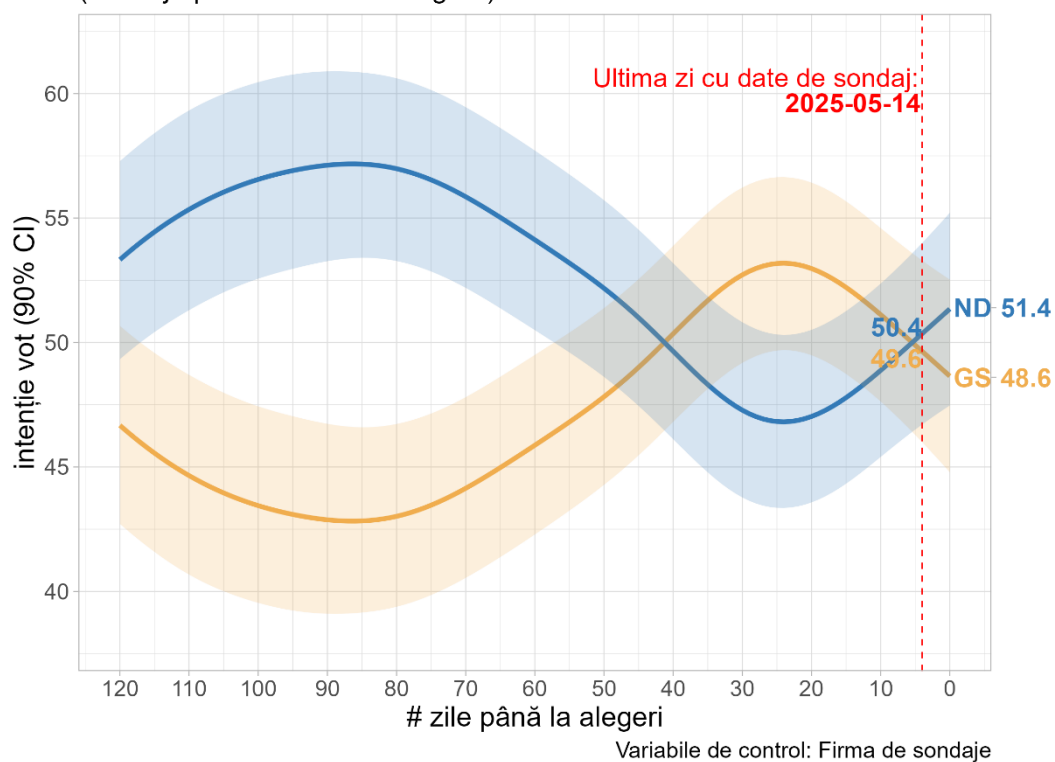
Prezidențiale 2025 (T2): Dinamica intențiilor de vot (sondaje publice - bayesian state-space model)



Prezidențiale 2025 (T2): Dinamica intențiilor de vot (sondaje publice - bayesian state-space model)



Prezidențiale 2025 (T2): Dinamica intențiilor de vot (sondaje publice - modelare gam)



Alegerile prezidențiale 2025, turul 2: o predicție a intențiilor de vot

Data: 2025-05-16

Cuvinte cheie: alegerile-prezidentiale-2025, modelarea-intențiilor-de-vot, predicții

În postările anterioare despre dinamica intențiilor de vot am analizat strict datele din sondajele publice cu scopul de a simplifica informația și a prezenta tendințele indicate de acestea. În alte postări am analizat variația estimărilor în funcție de diferiți factori (firma care a realizat sondajul, eventualele conflicte de interese, respectiv modalitatea de colectare a datelor) cu scopul de a arăta că estimările produse de sondaje sunt, în unele cazuri, puternic dependente de acești factori. Folosesc aici rezultatele acestor analize împreună cu alte date și analize pentru a produce o predicție a intențiilor de vot în cazul alegerilor de duminică.

Pentru realizarea predicției am combinat mai multe predicții, produse prin diferite metode, toate bazate pe datele din sondajele publice. De asemenea, am folosit rezultatele oficiale de la primul tur și două seturi de date brute. Predicția finală este o medie ponderată a cinci tipuri de predicții, majoritatea construite pornind de la estimările din sondajele publice. Cele cinci tipuri de predicții sunt următoarele:

- medii ale estimărilor realizate de firmele de sondare; am folosit doar sondajele cele mai recente, realizate de firme care au un grad relativ mai mare de încredere;

- o modelare statistică a dinamicii intențiilor de vot conform sondajelor publice (poll of polls; bayesian state-space model); intenția a fost să includ și modelul EuropeElects (a avut o performanță bună la T1), însă EuropeElects nu a produs un model și pentru T2;
- folosind două baze de date brute, am construit și evaluat o serie de variante de ponderări, asociate unor scenarii diferite; am folosit media a două dintre predicțiile obținute; principala limită a acestui demers o constituie faptul că două dintre baze nu acoperă populația în vârstă (65+); am încercat să corectez posibilul efect indus;
- folosind aceleași seturi de date brute, am realizat o modelare statistică de tip multi-nivel cu post-stratificare; desigur, ar fi fost de dorit ca toate bazele să acopere și populația vârstnică; o altă limită este numărul relativ redus de cazuri; și în acest caz am folosit media ponderată a mai multor predicții;
- pornind de la datele oficiale și de la estimări din sondaje, am aproximat distribuția așteptată a voturilor de la T2; de asemenea, am făcut estimări cu privire la numărul și structura celor care au votat la T1 și nu votează la T2, respectiv a celor care nu au votat la T1 și votează la T2.

Majoritatea sondajelor publice acoperă doar populația din România. Pentru a produce o predicție care să includă și diaspora, am folosit distribuția voturilor de la T1, estimări ale intențiilor de vot ale alegătorilor din diaspora conform datelor din sondajele AtlasIntel și estimările mele cu privire la împărțirea probabilă a voturilor.

Toate informațiile / estimările / aproximările cu care am lucrat au un anumit grad de incertitudine. Prin urmare, nu am folosit valori exacte ci distribuții asociate acestor valori. Am considerat că distribuțiile sunt de tip normal (formă de clopot). Prin faptul că folosește distribuții, modelarea ține cont de incertitudinile asociate valorilor de input. Desigur, distribuțiile au grade diferite de incertitudine. Le-am estimat folosind variațiile dintre estimările medii produse de cele patru metode menționate mai

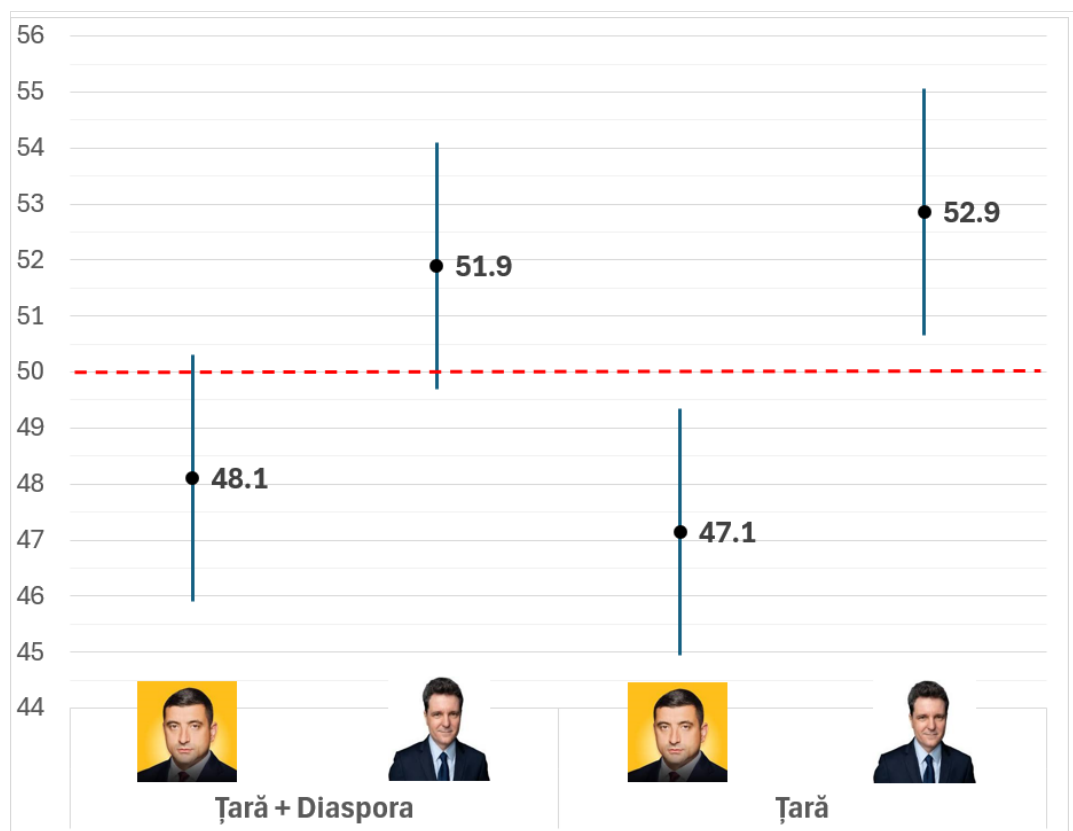
sus. Un model care nu ține cont de aceste aspecte riscă să producă estimări care (1) sunt prea sigure în relație cu informația disponibilă sau, dimpotrivă, (2) sunt prea imprecise, deci prea puțin informative sau (3) asociază același grad de incertitudine a estimării tuturor candidaților.

Foarte probabil, modelul de agregare a informațiilor descris pe scurt mai sus nu va produce estimări perfecte. Faptul că multe dintre informațiile folosite pentru realizarea acestei predicții provin din sondajele publice reprezintă o limită importantă. Spun asta mai ales pentru că majoritatea firmelor care au produs aceste sondaje nu au făcut publice informații esențiale. Astfel, nu știu dacă în realizarea estimărilor firmele au **(1) ținut sau nu cont de efectul de (in)dezirabilitate socială, (2) dacă și cum anume au ponderat datele brute și (3) dacă și cum anume au tratat indecisiile. Astfel de decizii metodologice pot influența estimările produse cu până la două-trei puncte procentuale (într-un sens sau altul).** Dacă aș fi avut aceste informații, aș fi putut să corectez mai bine eventualele erori sistematice (distorsiuni) asociate estimărilor din sondaje. Firmele de sondare care își propun să fie mai transparente cu privire la metodologie se pot inspira din modelul pe care l-am elaborat pentru MKOR și care este [public pe site-ul lor](#). Apreciez că această transparență ar trebui să caracterizeze toate firmelor de sondare a opiniei publice, mai ales în cazul sondajelor care prezintă estimări ale intențiilor de vot.

Pe scurt, am corectat estimările sondajelor folosind informații publice (extrem de limitate), informații deduse din informațiile publice și, nu în ultimul rând, ceea ce știu din diferite surse despre practicile firmelor de sondaje din România. În pofida acestor limite, mă aștept ca scorurile finale să fie în interiorul intervalelor de încredere. Cu cât un scor va fi mai aproape de centrul intervalului de încredere, cu atât predicția va fi mai bună.

Prezint rezultatul modelării într-o formă grafică simplă: cercurile pline indică predicțiile, iar liniile indică incertitudinea acestora, adică intervalul în care ne așteptăm, cu o probabilitate de 95%, să se situeze rezultatul final. În partea din dreapta a graficului apare predicția la nivel național (fără diaspora), iar în stânga predicția pentru național plus diaspora. Mă aștept ca în țară diferența să fie

de aproximativ 6 pp (puncte procentuale) în favoarea lui Nicușor Dan, iar în diaspora 12 pp în favoarea lui George Simion. Combinând aceste estimări obținem predicția rezultatului final al alegerilor: o diferență de aproximativ 4 pp în favoarea lui Nicușor Dan. **Cel mai probabil, următorul președinte al României va fi Nicușor Dan.**



Prezidențiale 2025, T2: dinamica ratei de prezență

Data: 2025-05-18

Cuvinte cheie: alegerile-prezidentiale-2025, rata-de-prezentat-la-vot

Un grafic cu rata de prezență la interval de 10 minute, alegerile prezidențiale 2025: T1 vs. T2. Voi schimba graficul aici, din când în când.

10:00 Comparativ cu T1, o rată de prezență mai mare cu 4 puncte procentuale (~ 14 pp față de 10). Pantă ceva mai abruptă față de T1 (tare nerăbdători oamenii). Ritmul de creștere se va reduce pe parcurs. Sunt șanse foarte mari să treacă de 60%. Poate ajunge chiar la 65%.

11:00 Crește panta!

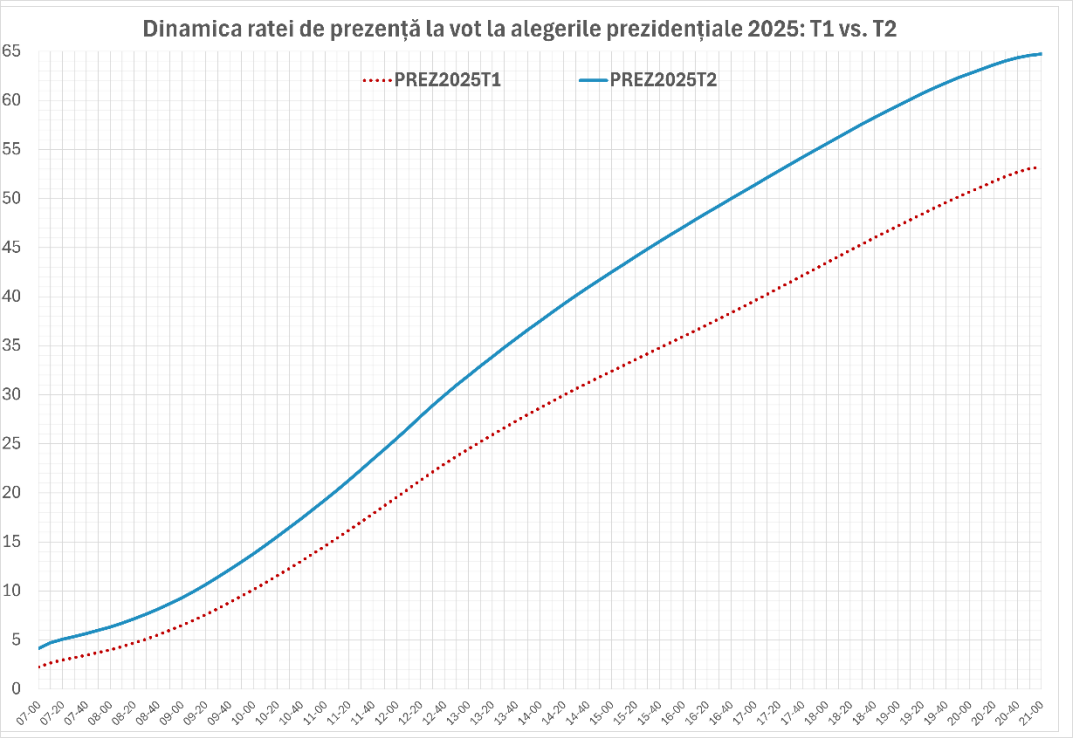
13:00 Crește și mai mult!

15:00 Crește și mai mult! +10 pp față de T1!

17:00 „Coafura rezistă!” În continuare panta o să scadă puțin câte puțin.

19:00 60% și mai avem două ore!

21:00 S-a oprit în intervalul pe care îl așteptam azi dimineață (60-65%).



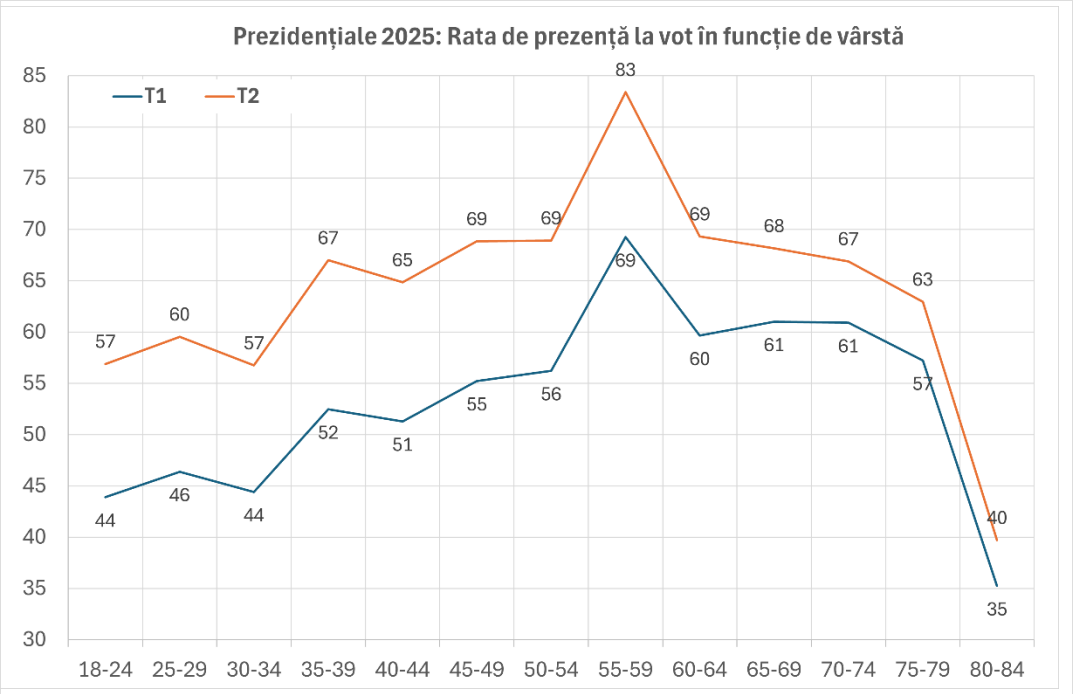
Prezidențiale 2025, T2: rata de prezență la vot în funcție de vârstă

Data: 2025-05-18

Cuvinte cheie: alegerile-prezidentiale-2025, rata-de-prezentat-la-vot

Pun aici o estimare a ratei de prezență la vot în funcție de vârstă, comparativ T1 vs. T2. Datele pentru T2 vor fi actualizate pe parcursul zilei. Distanța dintre cele două curbe indică deficitul/surplusul de votanți relativ la T1 în funcție de vârstă. Vârsta e legată de vot, nu la fel de intens precum educația formală, deci ...

Pentru estimarea ratelor am folosit date AEP (numărul votanților), INS (populația rezidentă) și Eurostat (numărul emigranților). Desigur, ar fi fost mai simplu dacă AEP ar fi făcut publică structura pe vârste a alegătorilor, nu a doar a votanților. Iar estimările din grafic ar fi fost mai apropiate de realitate. Nu mai reiau toate explicațiile. Cine dorește, poate citi postarea anterioară (prezidențiale 2024) pe aceeași temă [aici](#).



Calitatea estimărilor sondajelor: Prezidențiale 2025, T2

Data: 2025-05-19

Cuvinte cheie: alegerile-prezidentiale-2025, performanta-sondajelor

Cu ocazia alegerilor din toamna anului trecut am realizat două analize cu privire la calitatea estimărilor produse de sondaje (vezi [aici](#) și [aici](#)). Am făcut o analiză similară și pentru [turul 1](#). Astfel de analize sunt importante din multe puncte de vedere și pentru diferite tipuri de actori (media, publicul larg, institutele de sondaje, beneficiarii actuali și potențiali ai acestora).

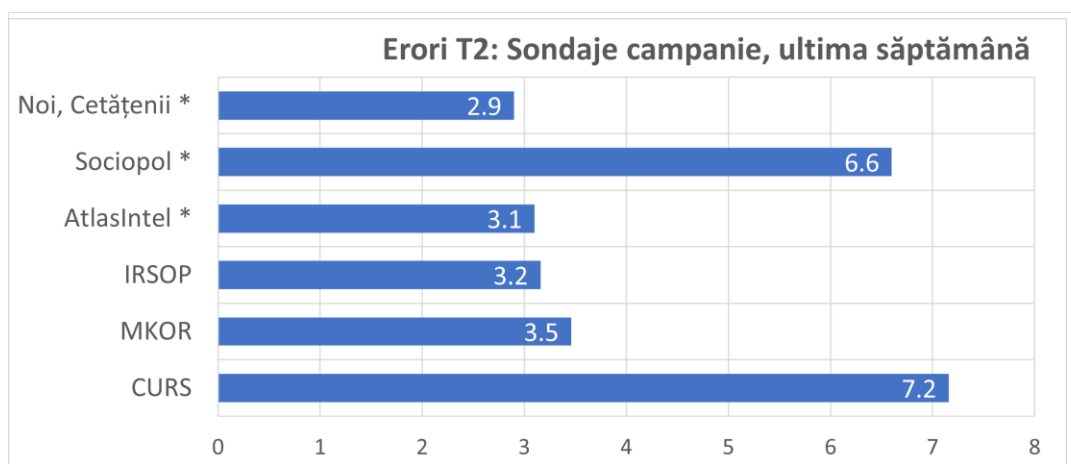
Dat fiind faptul că realitatea măsurată de sondaje este adesea fluidă, e important să distingem între sondajele realizate în săptămâna anterioară alegerilor și cele realizate în ziua alegerilor. Suplimentar, voi analiza și performanța unor modele de predicție bazate pe sondaje. Pentru a ține cont de aceste aspecte, am separat analiza în trei secțiuni: (1) sondaje realizate cu cel mult 7 zile înainte de data alegerilor, (2) modelarea estimărilor / datelor din sondaje și (3) sondaje realizate în ziua votului. Înainte de a trece la prezentarea rezultatelor analizei, sunt necesare câteva precizări metodologice, unele ceva mai tehnice.

Unele sondaje au fost realizate la nivel național (au acoperit doar populația din țară), prin urmare am comparat estimările produse de sondaje cu rezultatul final la nivel de țară. Situațiile care au la final semnul * indică faptul că sondajul / predicția a fost la nivel țară + diaspora, comparația fiind făcută cu rezultatul oficial de la același nivel.

Pentru a măsura calitatea estimărilor am folosit doar măsura MAE (fiind doar doi candidați, RMSE nu diferă). MAE (Mean Average Error) se referă la media erorilor absolute, eroarea fiind definită ca diferența absolută dintre procentul estimat de sondaj și procentul obținut conform BEC. O astfel de măsură indică distanța medie absolută dintre estimările produse de sondaje și rezultatele finale. O valoare mai mare indică o diferență (~ eroare) medie mai mare între estimări și rezultatul oficial. **Diferențele mai mari de trei puncte procentuale pot fi considerate problematice** (două puncte procentuale în cazul sondajelor realizate pe eșantioane mari - minimum 2000 „cazuri efective” = numărul cazurilor recalculat pentru situația în care eșantionul a fost de tip SRS (Simple Random Sample), adică simplu aleator.

Estimările sondajelor realizate în ultimele 7 zile

Cu două excepții, sondajele realizate cu cel mult 7 zile înainte de data alegerilor au avut erori medii apropiate de pragul de trei puncte procentuale (pp). Suplimentar, sondajul MKOR a indicat faptul că diferența dintre candidați crește în favoarea lui Nicușor Dan. Două sondaje, CURS și Sociopol, au avut diferențe semnificativ mai mari, duble comparativ cu pragul. Ultima zi de teren a sondajului CURS a fost cu 7 zile înainte de alegeri, de aici și cea mai mare parte a diferenței observate (în ultima săptămână dinamica intențiilor de vot a fost în favoarea lui Nicușor Dan). În cazul sondajului Sociopol (probabil o predicție bazată pe date de sondaj plus un model de alocare a voturilor din străinătate), [publicat pe FB](#) în ultima zi de campanie, cea mai mare parte a erorii nu mai poate fi explicată prin dinamica intențiilor de vot.



Medii și modelări ale estimărilor din sondaje

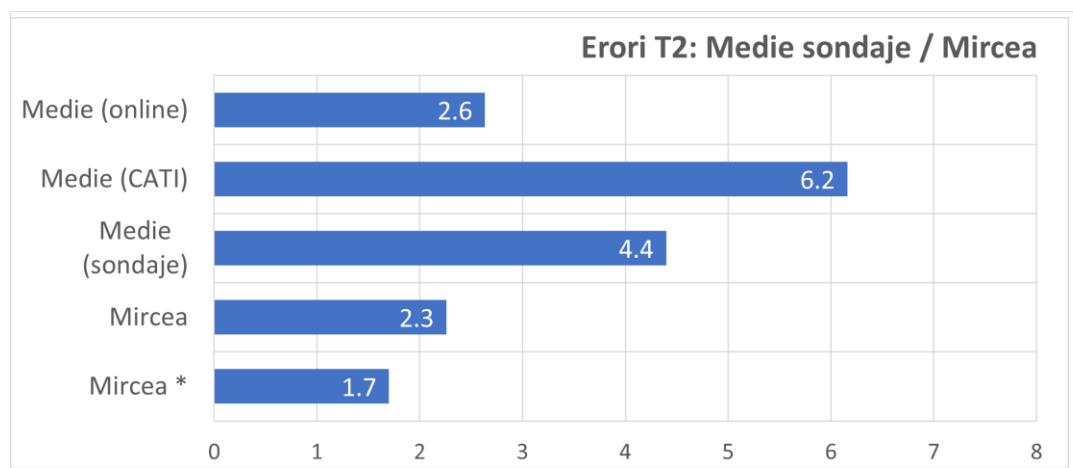
Ce se întâmplă dacă facem o medie a estimărilor din sondaje, dacă modelăm datele din sondaje, respectiv dacă realizăm un model care ia în calcul, pe lângă estimările din sondaje, și alte informații? Care dintre aceste variante produce predicții mai bune? Pentru a răspunde la această întrebare am procedat astfel:

- Am calculat media simplă a estimărilor produse de sondajele realizate cu cel mult 7 zile înainte de alegeri.
- Am calculat aceeași medie în cazul sondajelor CATI și a celor online. Cum unele sondaje au acoperit țara iar altele țara și diaspora, am comparat estimările obținute cu media estimărilor oficiale.
- Am inclus predicția realizată de mine și publicată pe blog în ziua de vineri (vezi [aici](#)). Am inclus ambele predicții (cu și fără voturile din diaspora).

O medie simplă a sondajelor din ultimele 7 zile (Medie sondaje) a produs o estimare care depășește pragul de 3 pp. Media erorilor diferă mult în funcție de tehnica de sondaj cu o mențiune importantă: ambele sondaje CATI au avut erori mari din cauze care, foarte probabil, nu au legătură cu metodologia CATI.

Predicțiile realizate de mine, prin agregarea și modelarea estimărilor din sondaje și a două baze brute, au avut cele mai mici erori. Predicția care a vizat rezultatul final (Mircea *), adică țară plus diaspora, a avut cele mai mici erori dintre toate estimările bazate pe date agregate (MAE = 1.7). Important, eroarea observată nu depășește marja de eroare indicată de predicție. Această predicție este semnificativ mai bună comparativ cu oricare dintre predicțiile făcute de sondajele pre-electorale. Simultan, predicția rezultatului final (Mircea *) a fost mai bună decât predicția sondajului Sociopol din ziua alegerilor și a avut o eroare doar puțin mai mare comparativ cu predicțiile CURS și Avangarde (exit-poll), respectiv Inscop (sondaj telefonic în ziua votului).

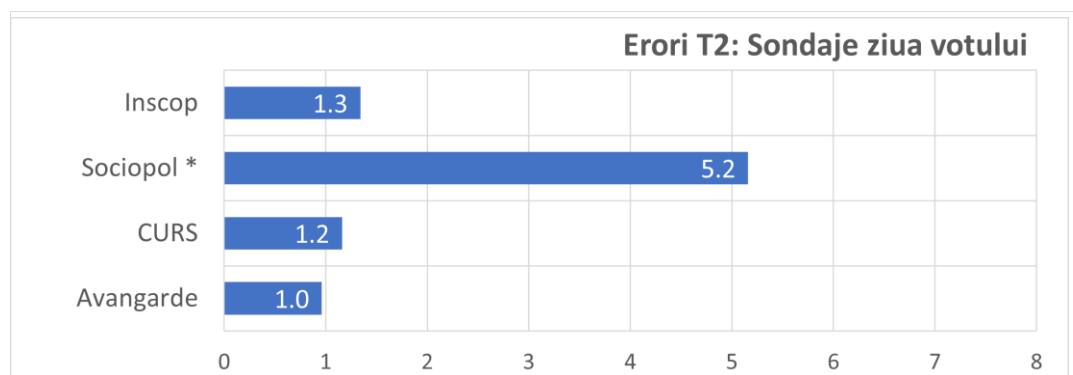
Predicția Mircea * are o performanță superioară predicției Mircea în principal ca urmare a faptului că am anticipat aproape perfect votul din diaspora: m-am așteptat la o diferență de 12 pp în favoarea lui George Simion și a fost o diferență de 11.7 pp.



Estimările sondajelor din ziua votului

Cu o singură excepție (Sociopol), sondajele din ziua votului au produs, evident, cele mai bune predicții. De această dată erorile în funcție de tehnica de sondaj (exit-poll vs. telefonic) au fost similare, semn că firmele care au realizat exit-poll-urile au

învăţat ceva din erorile de la T1. O ipoteză alternativă: acele cauze ale erorilor mai degrabă independente de cercetător au acţionat la T2 cu o intensitate mai redusă comparativ cu T1. O alegere între aceste ipoteze se poate face doar pe baza comparării datelor brute şi a modului în care au fost calculate estimările în cele două situaţii (T1 şi T2).



Notă: Estimările CURS, Avangarde şi Sociopol sunt cele finale, aşa cum au fost prezentate la televizor şi în presă în seara alegerilor. Estimarea Inscop a fost preluată dintr-o [postare pe FB](#) din seara alegerilor a directorului Inscop. Pentru analiză am folosit mijlocul intervalului prezentat: 56–57% Nicuşor Dan, deci 56.5%. A doua zi acesta a [prezentat pe FB](#) o altă estimare, uşor diferită: 55.8%.

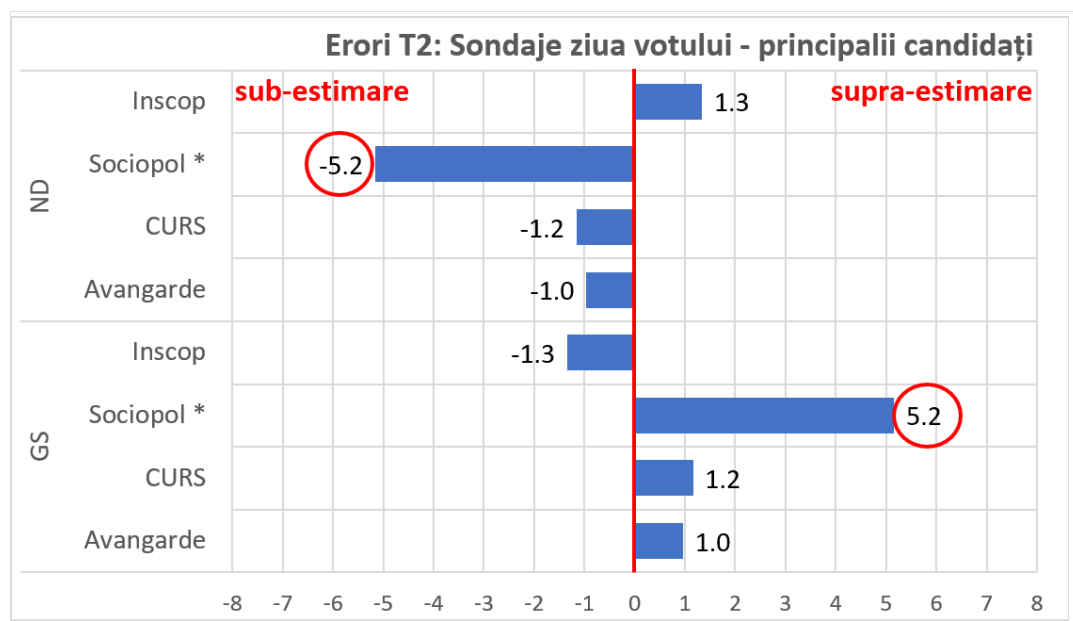
Sub/Supra-estimări în ziua votului

În graficele anterioare, distanţa dintre estimări şi realitate este calculată ca medie a unor diferenţe absolute. Pentru o imagine completă, e necesar să ne uităm şi la diferenţele punctuale observate. Procedând astfel, putem vedea şi sensul diferenţelor, respectiv putem face unele inferenţe cu privire la sursa erorilor. Astfel, putem observa dacă un candidat a fost sub-estimat sau supra-estimat în sondaje, respectiv dacă un eventual efect de acest tip variază sau nu în funcţie de compania care a realizat sondajul şi/sau tehnica de sondaj folosită. Prin sub-estimare mă refer la faptul că un candidat a obţinut în sondaje o estimare a intenţiei de vot mai mică faţă de procentul înregistrat la alegeri

(supra-estimarea se referă la fenomenul invers: un scor mai mare în sondaje comparativ cu scorul obținut la alegeri).

Pentru a nu lungi prea mult analiza, prezint doar evaluarea cu privire la sondajele realizate în ziua votului. Sociopol l-a supra-estimat mult pe George Simion și l-a sub-estimat mult pe Nicușor Dan. La celelalte firme diferențele au fost foarte mici: George Simion a fost supra-estimat puțin de CURS și Avangarde (probabil, pentru a evita eroarea de la T1 au „corectat” puțin mai mult decât era necesar), respectiv a fost sub-estimat puțin de Inscop.

Și o mică întrebare-mirare cu privire la exit-poll-ul Avangarde: Oare cum poate fi explicată scăderea de 0.7 pp a scorului lui Nicușor Dan în intervalul 19-21 în condițiile în care estimarea CURS a rămas practic neschimbată (+0.1), iar așteptarea ar fi fost ca Nicușor Dan să crească puțin în acel interval (la acele ore votează mai degrabă tinerii din urbanul mare)?!



Despre transparența sondajelor

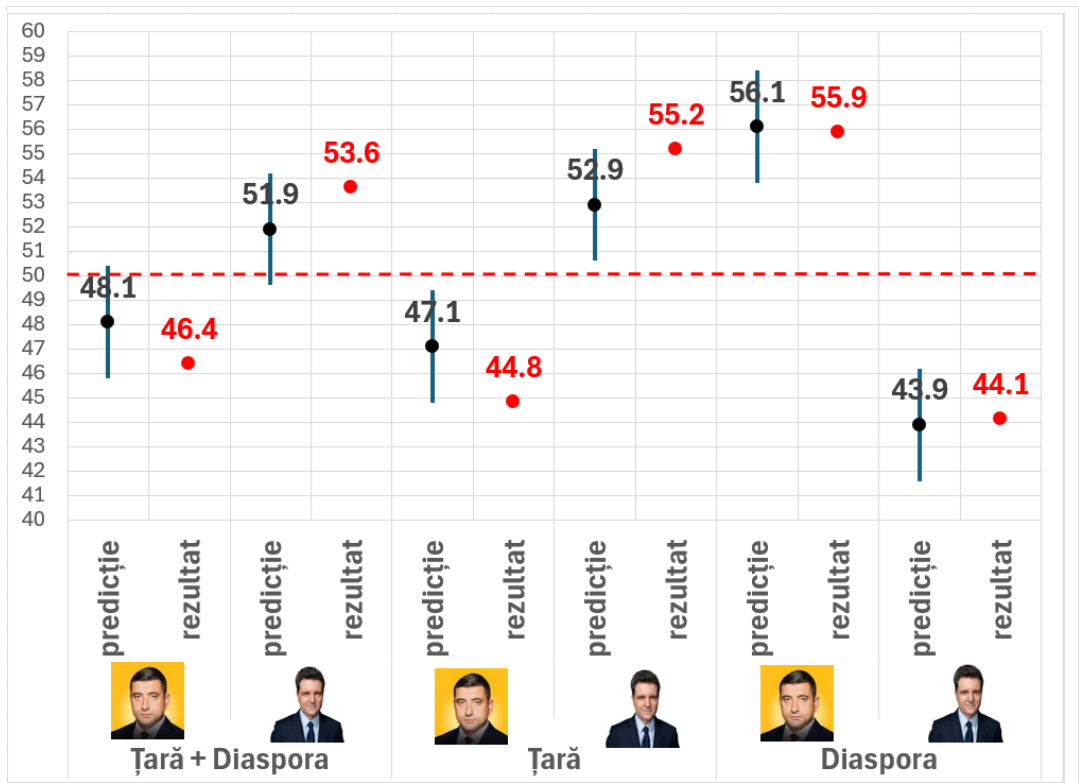
Un comentariu final, aparent marginal, dar extrem de important, cu privire la **transparența sondajelor din România**. Transparența este un criteriu important în funcție de care putem evalua calitatea unui sondaj și profesionalismul unui institut de sondaje. Prin transparență mă refer la tipul de informații care sunt făcute publice, alături de selecția de grafice, și la gradul de detaliere al acestor informații. În cazul sondajelor politice din România, cel mai adesea, sunt prezentate doar informații minimale, fără prea multe detalii și fără să fie acoperite o serie de informații care sunt foarte importante pentru cineva care dorește să analizeze calitatea datelor produse. Puține institute publică un raport pe propriul site, deci analizele pot fi făcute doar pe baza informațiilor care apar în media. În cazul sondajelor din ziua votului, lucrurile stau chiar și mai rău – până la acest moment, niciuna dintre firme nu a publicat un raport, fie el și minimal.

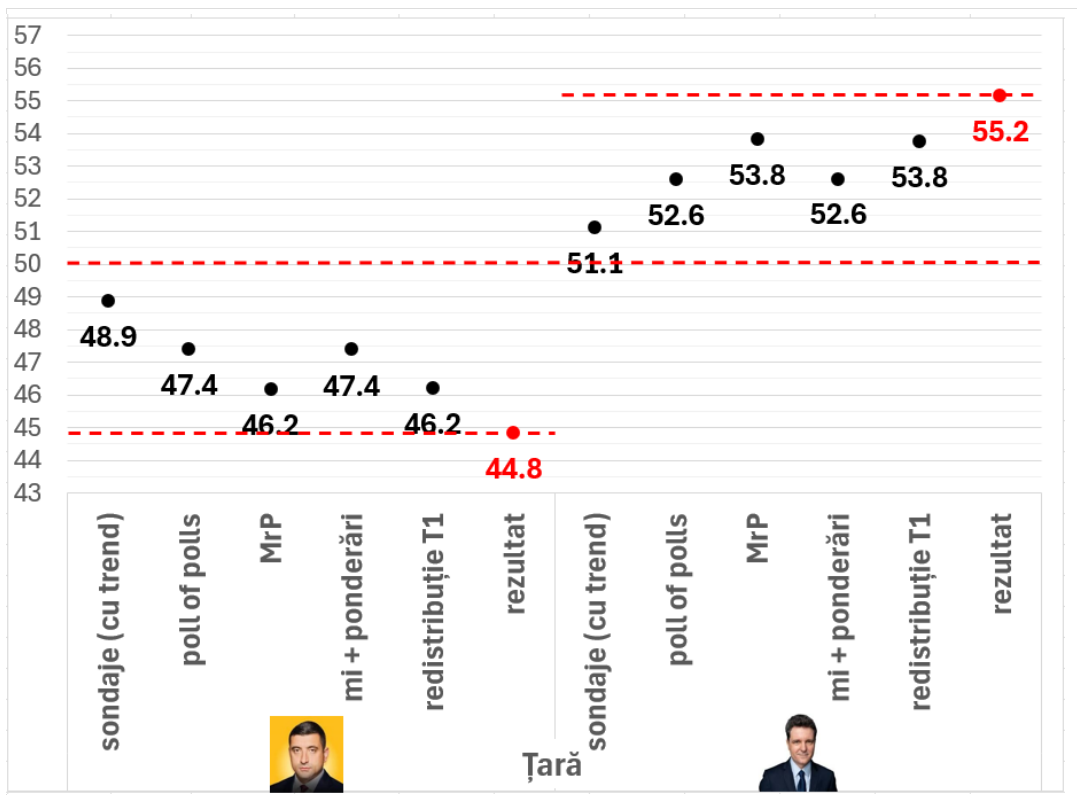
În cazul sondajelor pre-electorale apar informații minimale în presă în cazul CURS, IRSOP AtlasIntel și [Noi, Cetățenii](#). Din nou, singura firmă care a publicat un [raport extins](#), atât metodologic cât și de analiză a rezultatelor, a fost MKOR. Pentru a surprinde dinamica intențiilor de vot din campanie, în sondajul MKOR am folosit un design longitudinal de tip trend. Rezultatele prezentate în raport au arătat că opțiunile de vot nu sunt statice, că unul dintre candidați (Nicușor Dan) a trecut în frunte și că, foarte probabil, distanța dintre candidați va crește până în ziua votului. Raportul a indicat modul de tratare a indecișilor (fie ei falși sau reali): aceștia au fost alocați unuia dintre candidați sau niciunuia folosind un model de predicție. Cei interesați pot citi raportul integral [aici](#). Pentru transparență, precizez următoarele: (1) MKOR este o companie relativ nouă, axată pe cercetarea de marketing; (2) populația măsurată în sondaj a fost 18-64 ani, preponderent din urban (publicul mai degrabă comercial), deci nu a acoperit toată populația cu drept de vot; (3) sondajul a fost de tip online; (4) la cererea lor, le-am oferit consultanță neplătită, strict pentru sondajele politice; (5) toate sondajele politice realizate și prezentate de MKOR în campanie au fost auto-finanțate.

Două grafice cât zece mii de cuvinte

Data: 2025-05-20

Cuvinte cheie: alegerile-prezidentiale-2025, calitatea-predictiilor, predictii





Drumul de la T1 la T2

Data: 2025-05-21

Cuvinte cheie: alegerile-prezidentiale-2025, transferul-de-voturi

Context: La T1 al alegerilor prezidențiale 2025, GS a obținut un număr aproape dublu de voturi comparativ cu ND. În proximitatea T1, cu referire la intenția de vot pentru T2, sondajele indicau un avans de aproximativ 15 puncte procentuale pentru GS. Cu toate acestea, ND a câștigat alegerile la o diferență de 7 puncte procentuale.

Întrebare: Cum se explică rezultatul „neașteptat” de la T2 sau, spus altfel, de unde au venit voturile care i-au adus lui ND victoria în alegeri?

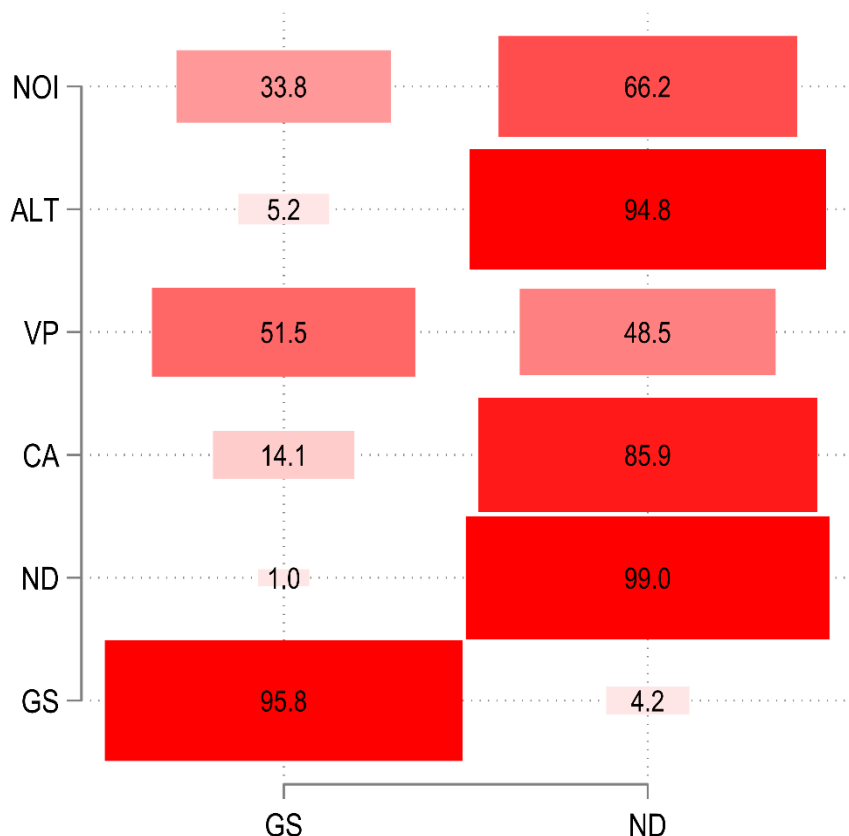
Metodologie: În extrem de puține situații institutele de sondaje din România fac publice bazele de date. Rapoartele publicate de aceleași institute prezintă rar tabelul de asociere dintre votul la T1 și T2. Chiar dacă aș fi avut acces la astfel de date, le-aș fi privit cu anumite rezerve, pentru motive ce țin de limite metodologice și practice. În principal este vorba despre numărul redus de cazuri asociat fiecărei perechi T1-T2. De exemplu, în cazul unui eșantion de 1000 respondenți, rămânem cu aproximativ 900 alegători care declară că vor vota / au votat, respectiv cu aproximativ 700 de alegători care declară intenția de a vota / votul efectiv pentru T1. Pentru situația în care avem cinci candidați mari la T1 (plus categoria celor care nu votează și la T2) și doi candidați la T2 (plus votanții noi), rezultă un număr de 18 perechi (T1 - T2), deci o medie de 39 cazuri pe pereche. O estimare care se bazează pe un număr atât de mic de cazuri are o imprecizie foarte mare. Mai

mult, în cazul unor perechi „sensibile,” tentația respondenților de a ascunde răspunsul sau de a declara un alt răspuns decât cel real este mai mare (de aici și erori de estimare mai mari). Simplu spus, unii alegători nu vor recunoaște că au votat / intenționează să voteze pentru un anumit candidat. În concluzie, o analiză la nivel individual, bazată pe declarații din sondaje, nu este cea mai bună strategie.

O soluție alternativă este analiza / inferența ecologică. E vorba de o analiză realizată pe date la nivel agregat cu scopul de a produce inferențe cu privire la situația de la nivel individual. În domeniul studiilor electorale analiza este cunoscută sub numele de „voter transition matrix / probabilities.” Pentru a reduce cât mai mult posibilitatea realizării unor inferențe incorecte, e important să folosim date situate la un nivel de agregare cât mai redus. Prin urmare, toate analizele asociate acestui text sunt realizate pe datele oficiale la nivel de secție de vot. Am considerat toate secțiile de vot, inclusiv cele din străinătate. Pentru a simplifica prezentarea datelor, am redus analiza la candidații care au obținut un număr relativ mai mare de voturi la T1.

Rezultate: Valorile din grafic sunt procente pe linii (totalul pe fiecare linie e 100%), deci indică, pentru fiecare candidat de la T1, ponderea voturilor care s-au dus spre fiecare dintre finaliști. La nivelul fiecărei perechi (celule / combinații T1+ T2), mărimea relativă a transferului este comunicată vizual prin intermediul intensității culorii și a ariei dreptunghiului. O culoare roșie mai intensă, respectiv un dreptunghi mai mare, indică o pondere mai mare a voturilor transferate.

Transferul de voturi de la T1 la T2 - TOTAL -



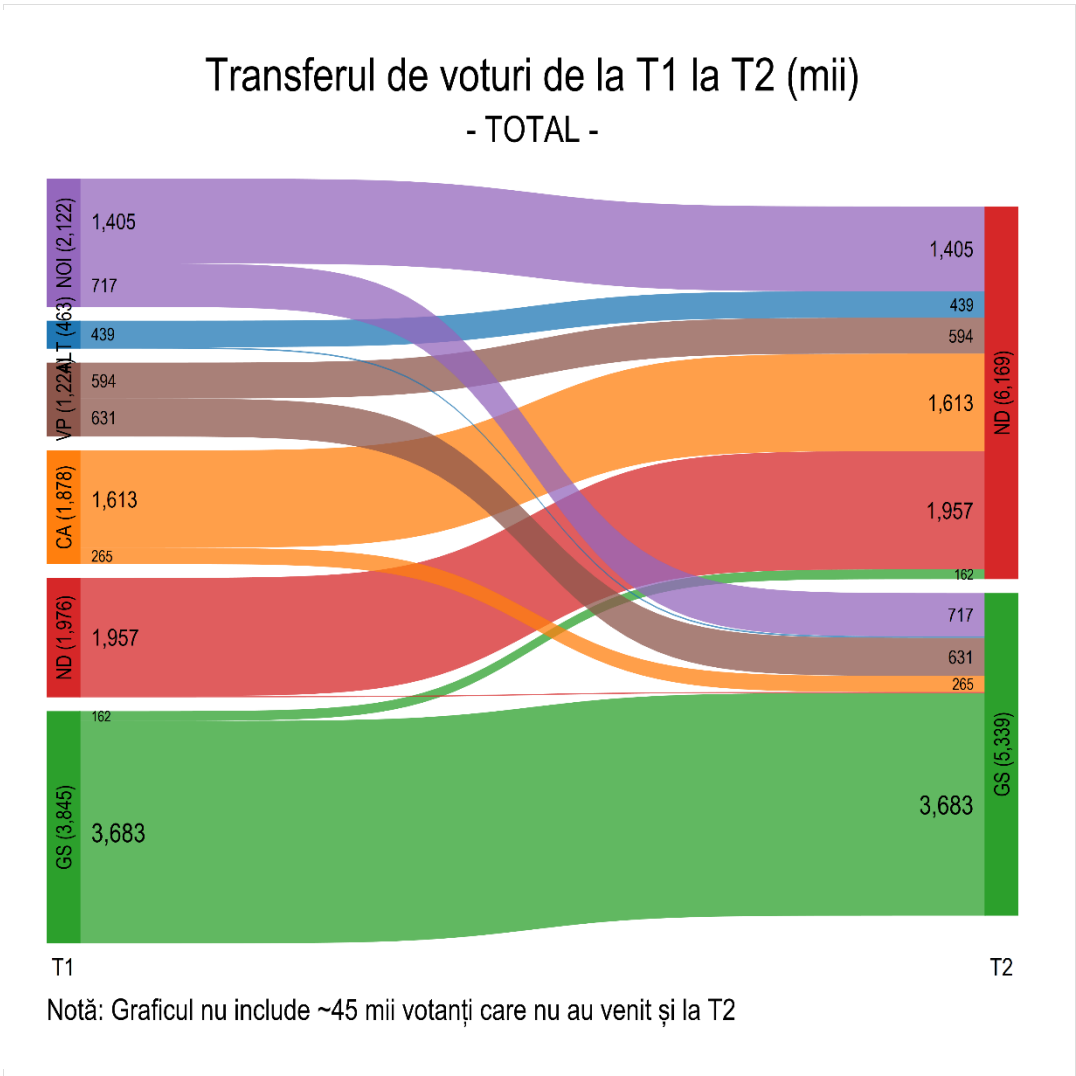
Notă: Graficul nu include ~45 mii votanți care nu au venit și la T2

Concluzii: Finaliștii și-au conservat cea mai mare parte a voturilor primite la T1. Comparativ cu GS, ND a reușit să păstreze o pondere mai mare a voturilor de la T1 (99% vs. 96%). ND a primit 86% din voturile lui CA, 49% din voturile lui VP și 95% din voturile celorlalți candidați. Foarte important, ND a beneficiat de 66% dintre voturile celor care s-au prezentat doar la T2.

Un grafic cu valorile absolute: Pun mai jos un grafic alternativ de ilustrare a transferului de voturi, de această dată folosind valori absolute (mii voturi). Pentru că am exclus categoria celor care au votat doar la T1, totalurile nu reproduc exact valorile oficiale.

GS și-a păstrat 3.7 milioane din zestrea de voturi de la T1 și a pierdut aproximativ 162 mii în favoarea lui ND. De asemenea, GS a beneficiat de 631 mii voturi de la VP și 265 mii de la CA. La acestea s-au adăugat 717 mii voturi de la cei care au venit doar la T2.

ND și-a păstrat aproape toată zestrea inițială de voturi și a beneficiat de 1.6 milioane voturi de la CA, 594 mii de la VP, 162 mii de la GS și 439 mii de la alți candidați. La aceste voturi s-au adăugat 1.4 milioane voturi de la cei care au venit doar la T2.

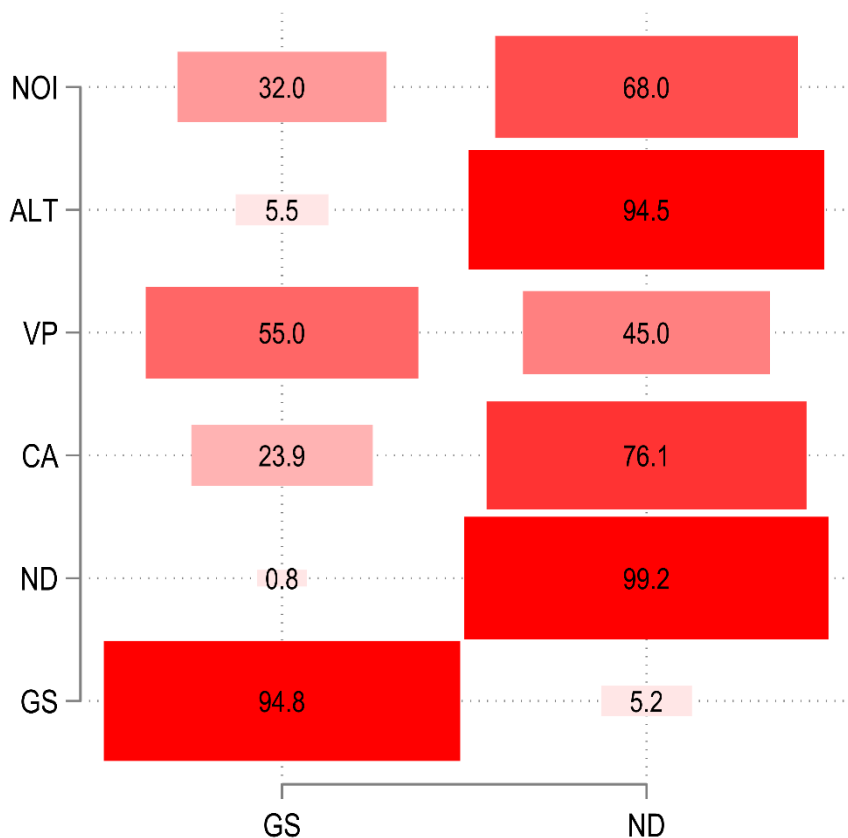


Transferul de voturi de la T1 la T2 - URBAN -

Categorie	GS	ND
NOI	26.0	74.0
ALT	5.9	94.1
VP	47.8	52.2
CA	5.6	94.4
ND	1.2	98.8
GS	95.4	4.6

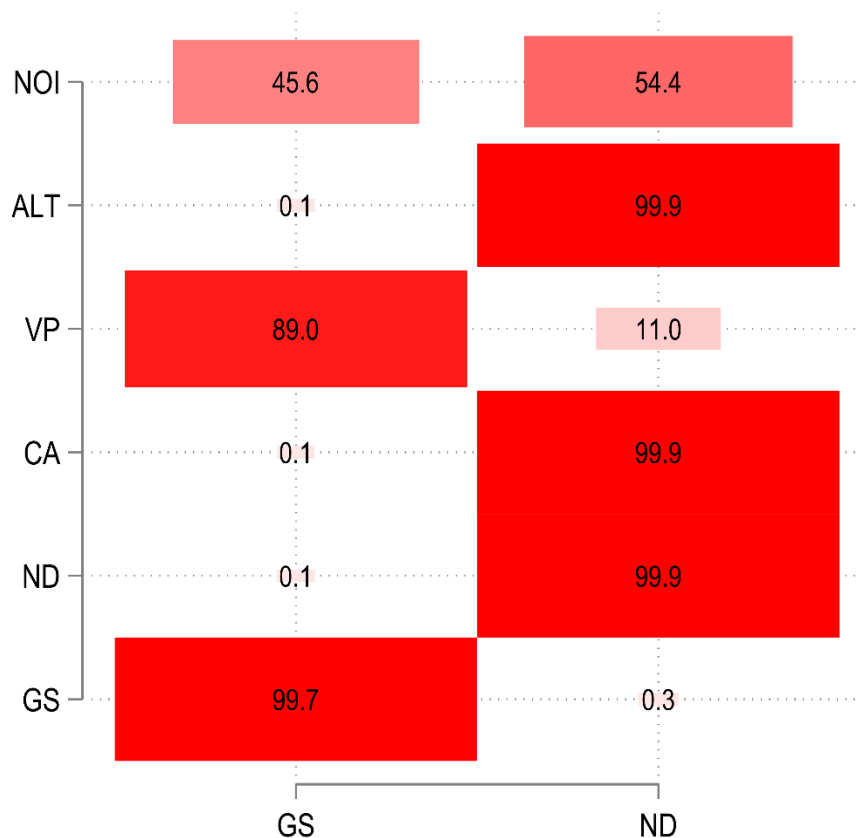
Notă: Graficul nu include votanții din urban care nu au venit și la T2

Transferul de voturi de la T1 la T2 - RURAL -



Notă: Graficul nu include votanții din rural care nu au venit și la T2

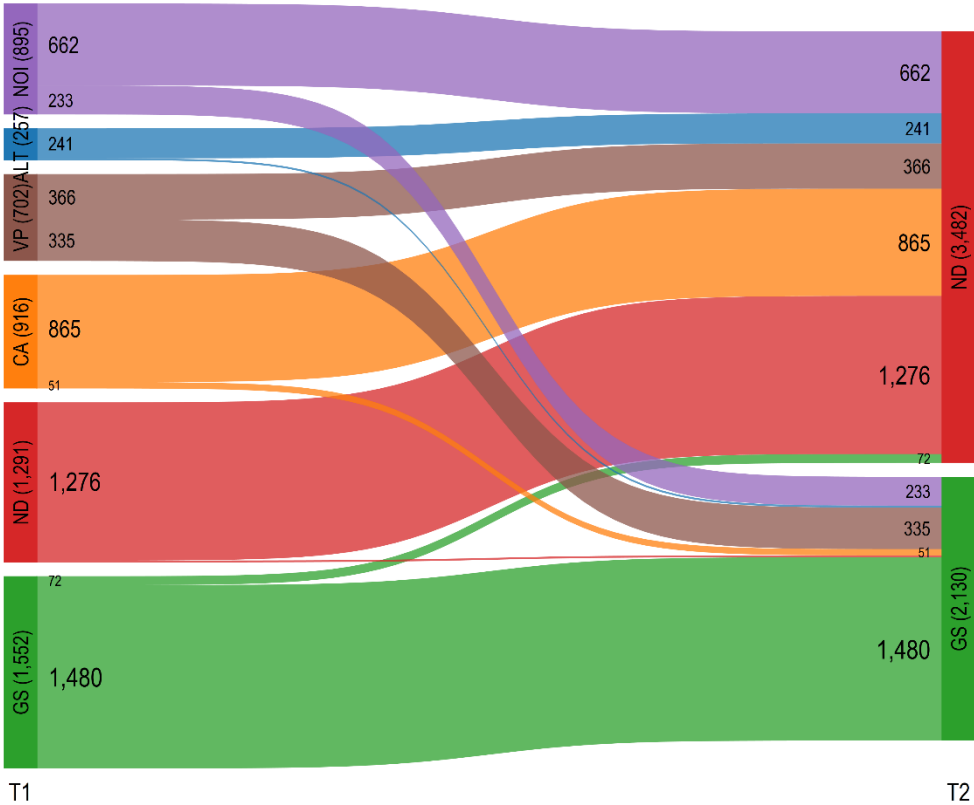
Transferul de voturi de la T1 la T2 - DIASPORA -



Notă: Graficul nu include votanții din diaspora care nu au venit și la T2

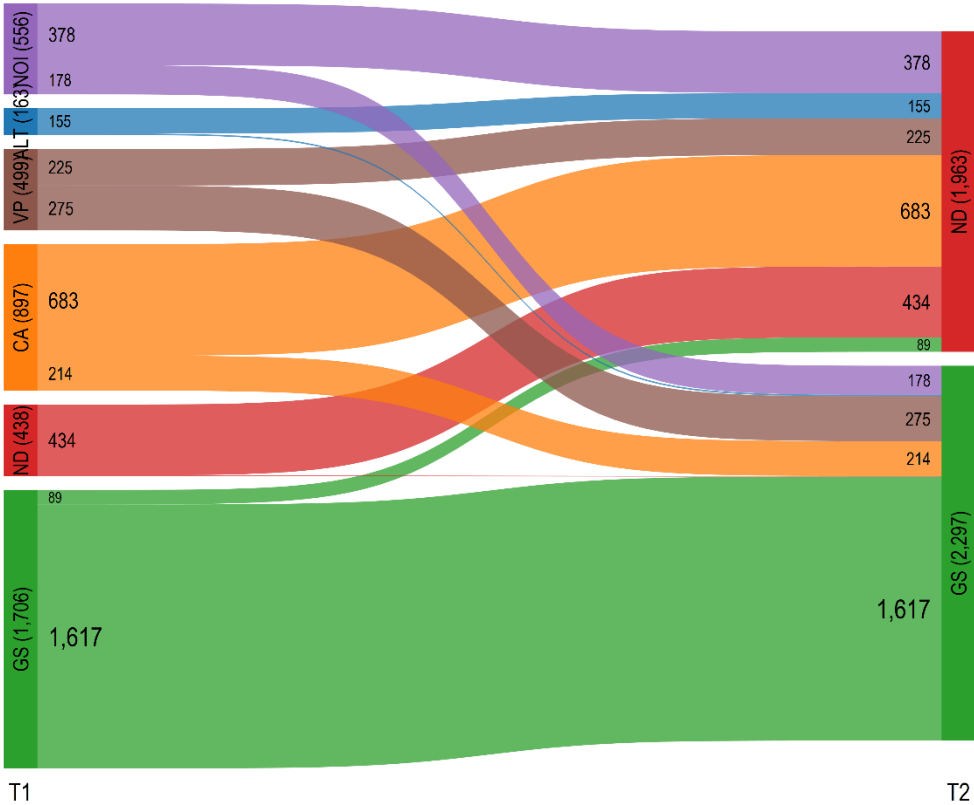
Adaug și graficele cu transferul voturilor (valori absolute; mii), separat în funcție de urban, rural și diaspora. Erau făcute de la început, dar nu am vrut să lungesc povestea. Sper să fie utile pentru cât mai mulți.

Transferul de voturi de la T1 la T2 (mii) - URBAN -



Graficul nu include votanții din urban care nu au venit și la T2

Transferul de voturi de la T1 la T2 (mii) - RURAL -



Graficul nu include votanții din rural care nu au venit și la T2

Transferul de voturi de la T1 la T2 (mii) - DIASPORA -



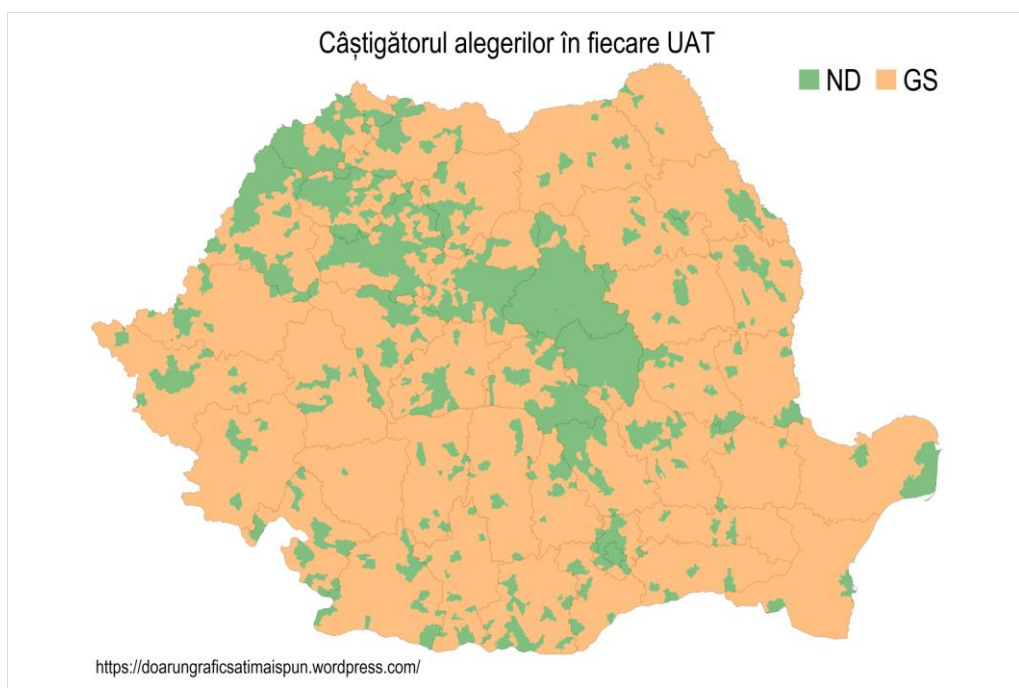
Graficul nu include votanții din diaspora care nu au venit și la T2

Land doesn't vote

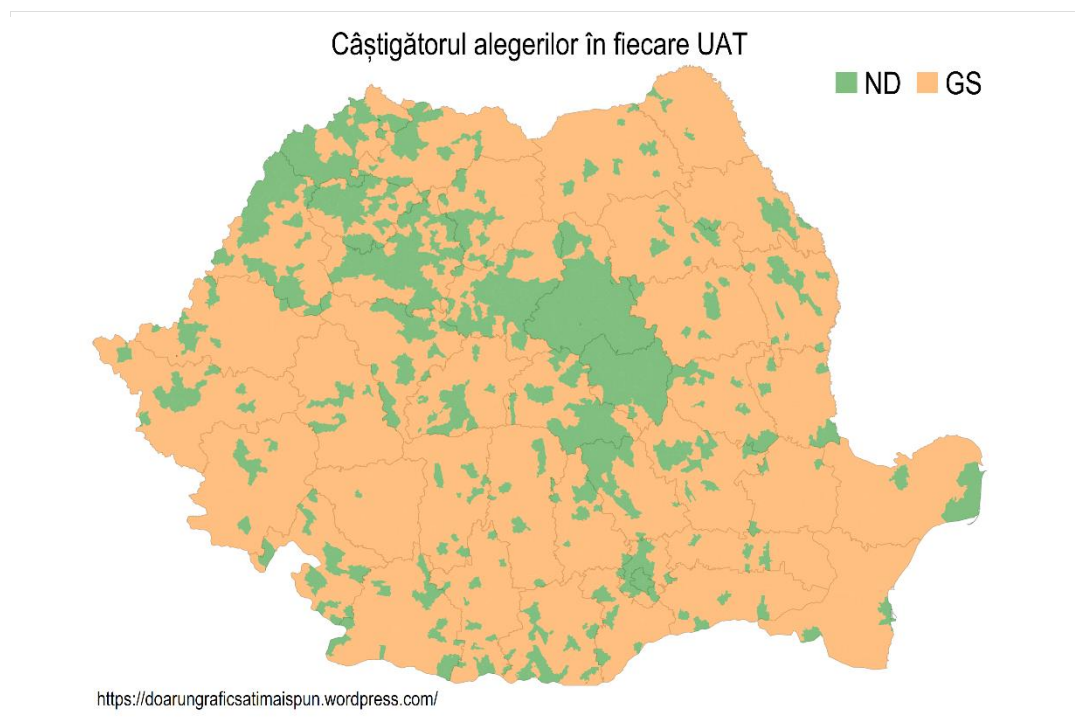
Data: 2025-05-27

Cuvinte cheie: alegerile-prezidentiale-2025, harti

La prima vedere titlul poate părea ciudat. De fapt, titlul complet ar fi trebuit să fie ceva de genul „Land doesn't vote. People do.” O reprezentare de acest tip a rezultatului alegerilor este destul de comună în cazul SUA (vezi un exemplu [aici](#)). Mesajul transmis este cât se poate de clar: reprezentările spațiale ale rezultatului alegerilor pot fi înșelătoare uneori. În această postare încerc să reprezint într-o manieră similară rezultatul alegerilor prezidențiale de săptămâna trecută. Iată ce a ieșit!



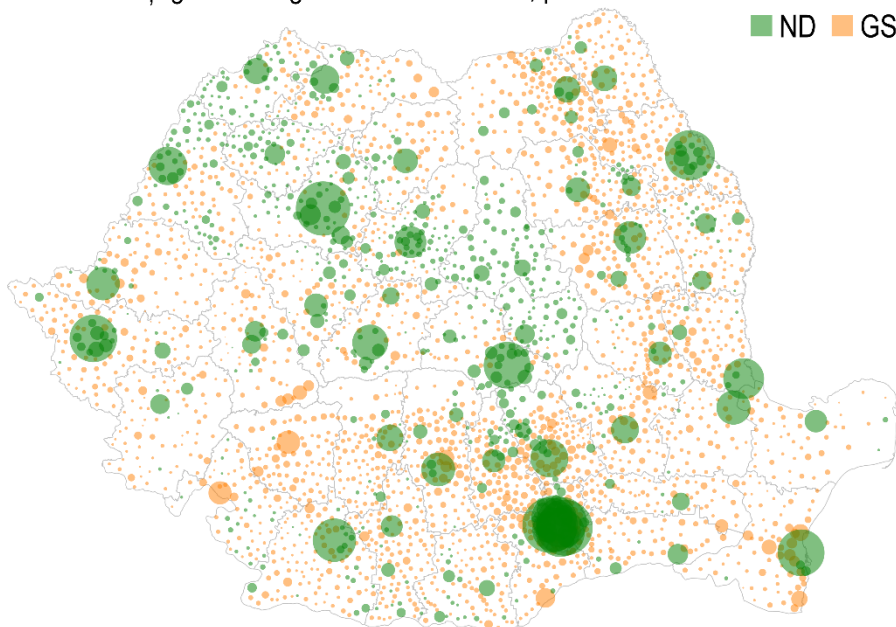
Și câteva explicații: Deseori, modul clasic de prezentare spațială a rezultatelor alegerilor poate fi înșelător. Să presupunem că nu știm rezultatul alegerilor și privim harta de mai jos. Oricine face acest lucru va crede că alegerile au fost câștigate detașat de George Simion.



Să privim acum harta cu aceleași rezultate, considerând în același timp și numărul absolut al voturilor primite de candidați. De această dată nu mai suntem atât de siguri cine a câștigat sau chiar percepem o victorie a candidatului Nicușor Dan.

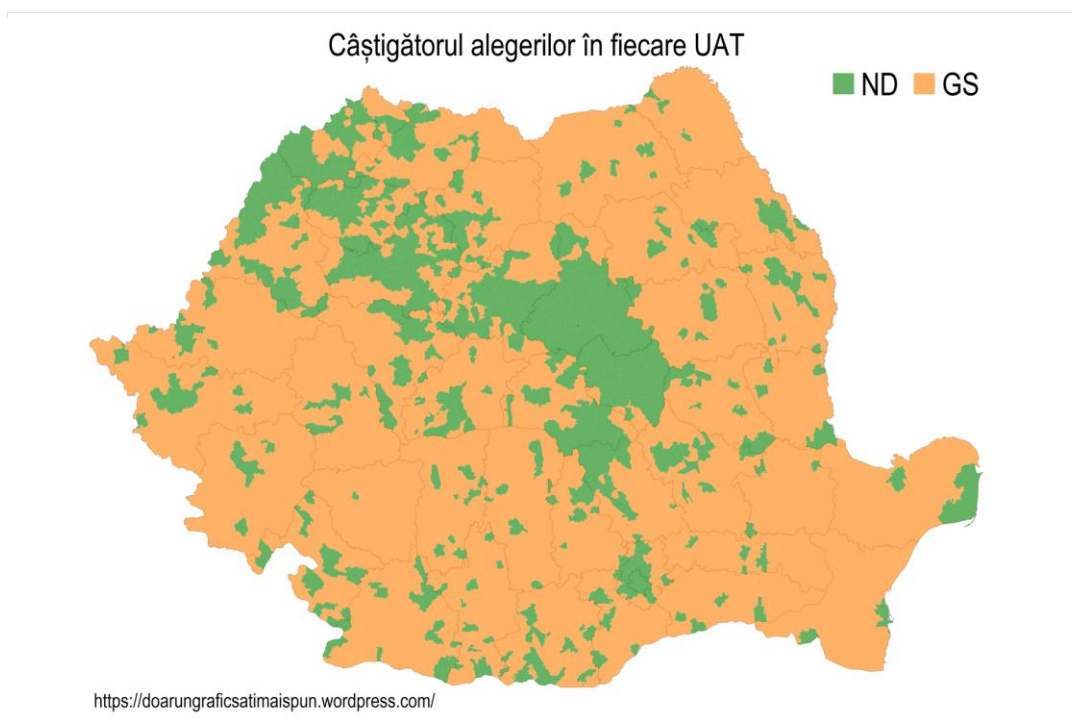
Câștigătorul alegerilor în fiecare UAT, ponderat cu # voturilor

■ ND ■ GS

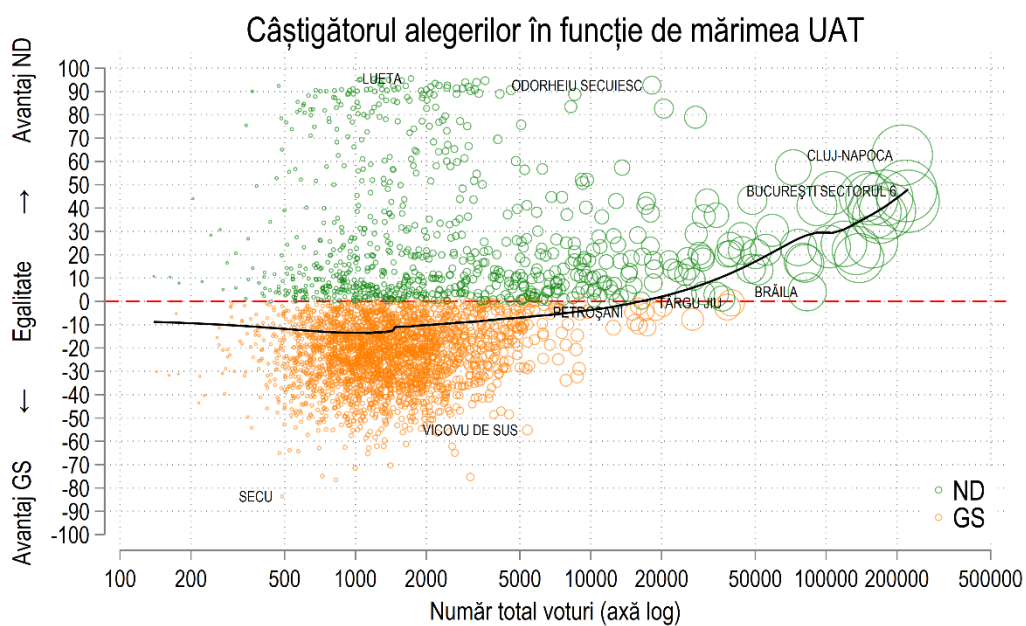


<https://doarungraficsatimaispun.wordpress.com/>

Notă: Ideal ar fi fost să modific automat poziția cercurilor astfel încât să nu apară suprapuneri. Cum pachetul utilizat nu permite acest lucru, am folosit transparența (în cazul capitalei ajută prea puțin). Pun mai jos o variantă de reprezentare în care localitățile sunt reprezentate prin aria acoperită (mai exact triplul acesteia, pentru a face vizibile și localitățile mici).



Pun mai jos un grafic care prezintă la nivel de UAT relația dintre marja victoriei și numărul total de voturi (atenție, axa X este reprezentată grafic pe o scală logaritmică). Câteva concluzii: pe măsură ce crește mărimea localității, șansele ca Nicușor Dan să câștige (detașat) alegerile cresc; această relația nu este liniară, iar creșterea are loc doar după ce localitatea depășește aproximativ două mii de voturi; Nicușor Dan a câștigat majoritatea localităților mari; George Simion a câștigat semnificativ mai multe localități mici; Nicușor Dan a câștigat detașat, adică la o diferență de cel puțin 80 puncte procentuale, relativ mai multe localități (majoritatea mici); raportat la localitățile relativ mai mari, cea mai strânsă competiție a fost în Petroșani, Târgu-Jiu și Brăila; Nicușor Dan a obținut cel mai mare scor în Lueta (Harghita), iar George Simion în Secu (Dolj); raportat la localitățile foarte mari, diferența cea mai mare a fost înregistrată în Cluj-Napoca.



<https://doarungraficsatimaispun.wordpress.com/>

Mărimea cercurilor indică numărul total de voturi

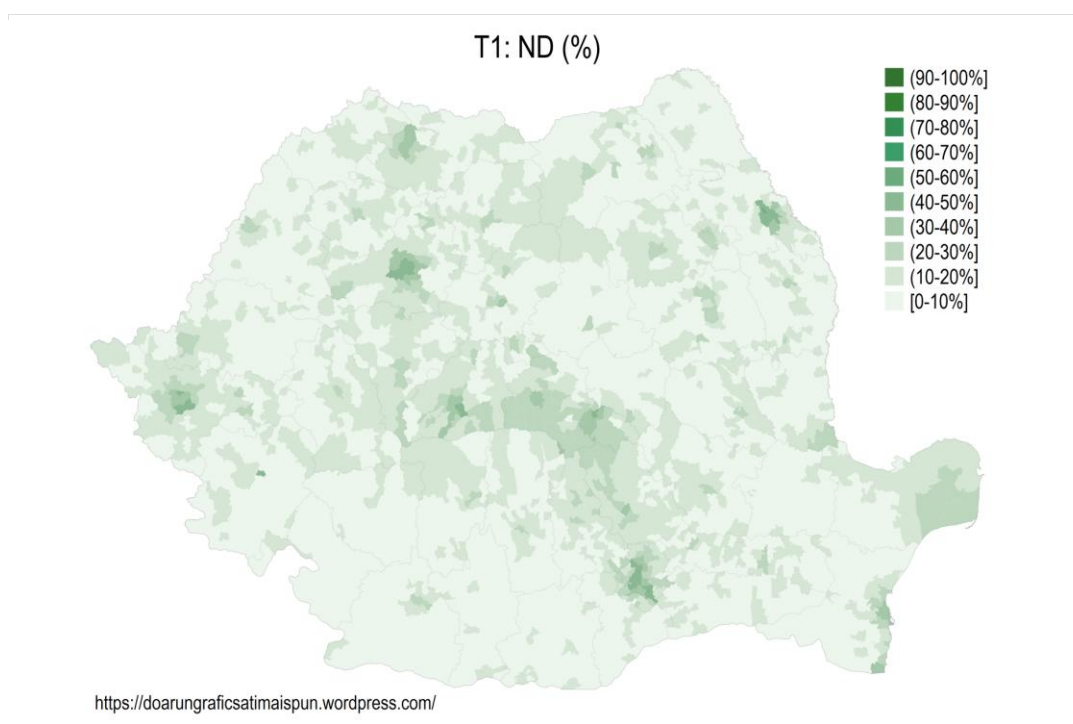
Schimbarea preferințelor de vot la nivel de UAT

Data: 2025-05-29

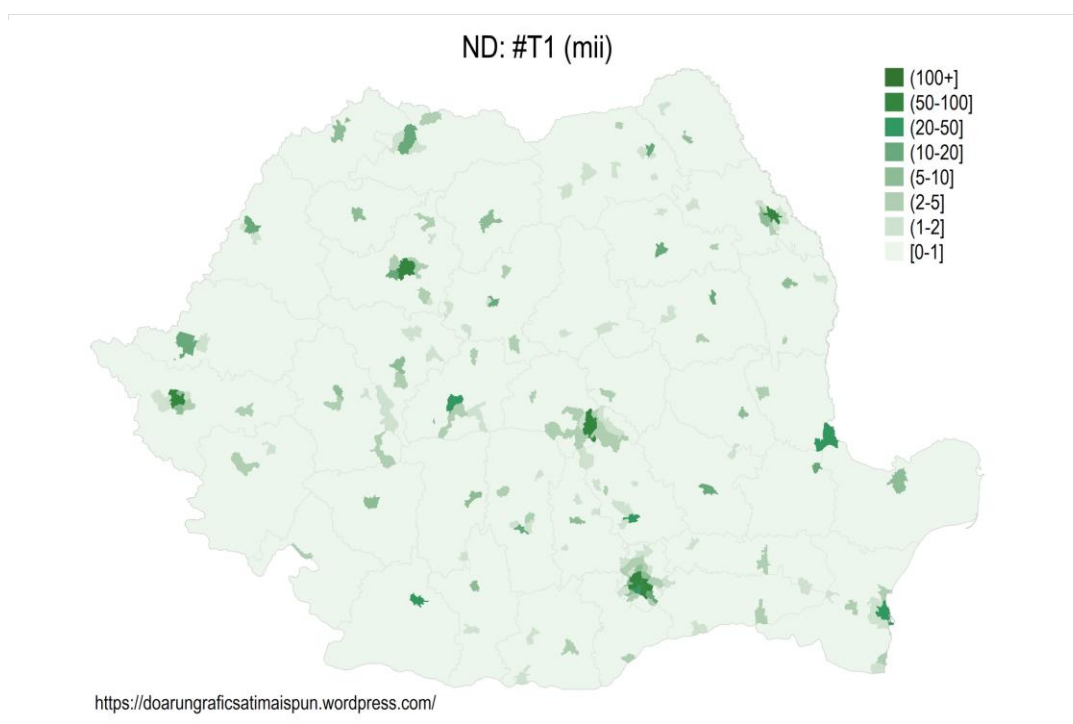
Cuvinte cheie: alegerile-prezidentiale-2025

Prezint aici câteva grafice dinamice care surprind schimbarea preferințelor de vot între T1 și T2 în profil teritorial. Fiecare grafic își propune să răspundă la o întrebare specifică.

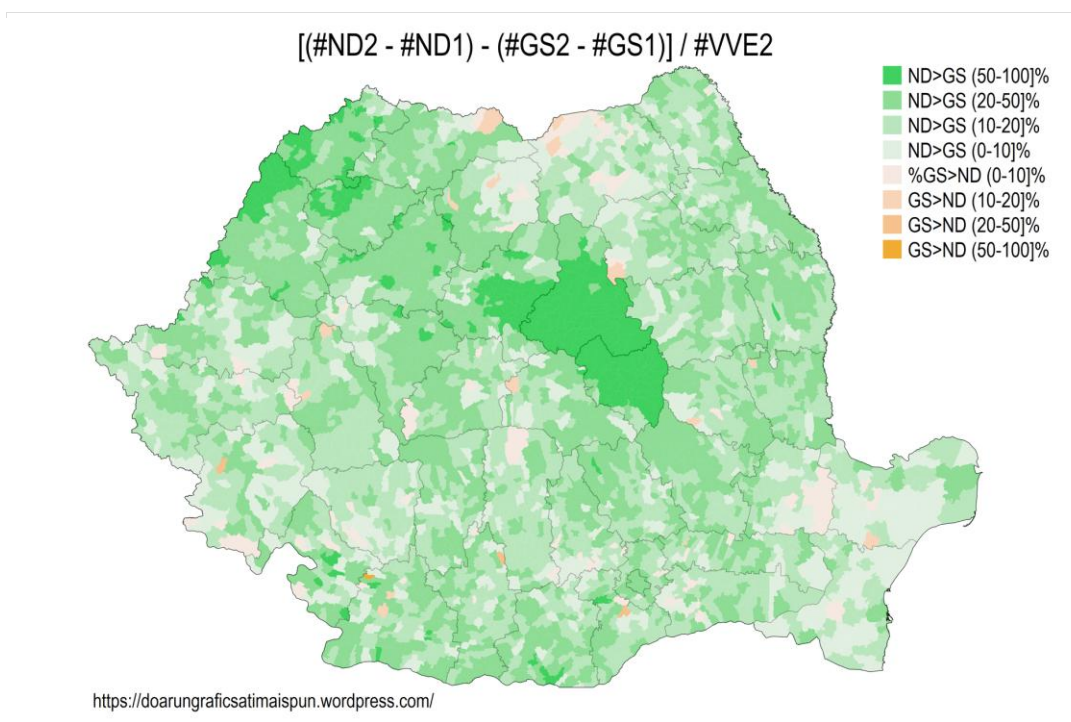
Cum s-a schimbat procentul voturilor obținute de ND și unde a fost înregistrată o creștere relativă mai mare? Dacă inițial ND a avut scoruri mai degrabă mici în majoritatea localităților, la T2 și-a dublat, triplat, cvadruplat ... procentele obținute la T1. Creșterile cele mai spectaculoase, în puncte procentuale, au fost în localitățile cu populație de etnie maghiară. Creșteri mari au fost și în localități din sud, sud-vest și est.



Cum s-a schimbat numărul voturilor obținute de ND și unde a fost înregistrat un plus mai mare de voturi? La T1 voturile pentru ND au fost preponderent în orașele foarte mari. Acestora, la T2, s-au alăturat și localitățile din proximitate. Cele mai mari creșteri în termeni absoluți au fost în orașelor foarte mari și localitățile din proximitatea acestora.

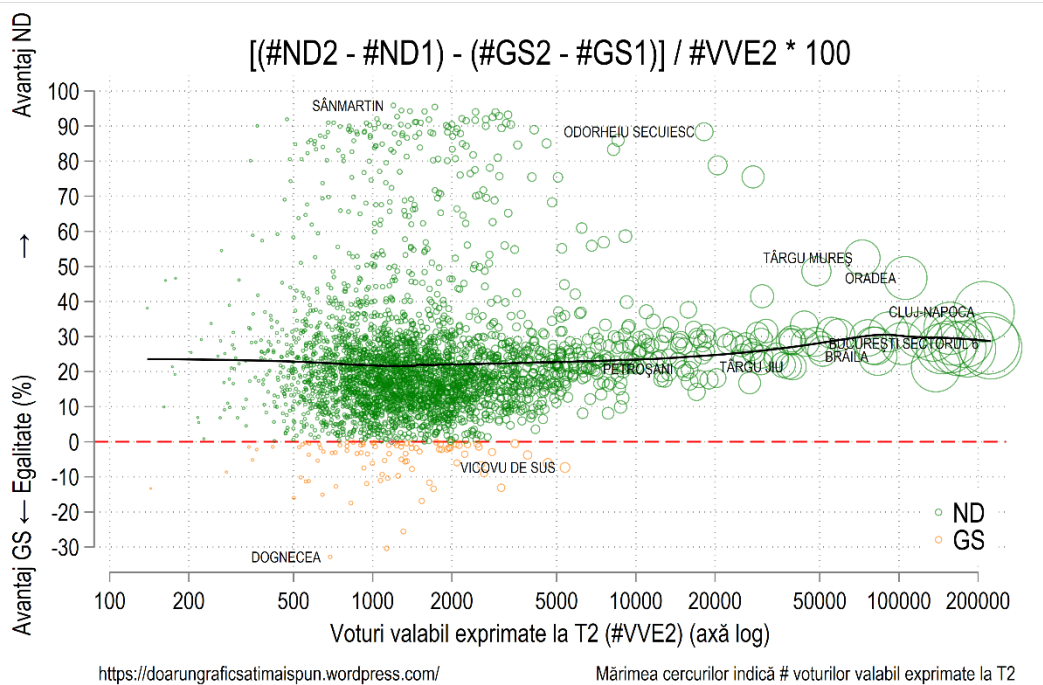


Ambii finaliști au obținut mai multe voturi la T2 comparativ cu T1. În cazul lui ND creșterea a fost semnificativ mai mare. Care a fost diferența dintre creșterile înregistrate de cei doi finaliști? În majoritatea cazurilor dinamica opțiunilor de vot a fost în favoarea lui ND. Preferința pentru ND observată în cazul orașelor (foarte) mari a adus, firesc, și cele mai multe voturi.



Continuând analiza anterioară, ne putem întreba dacă diferența între creșterea numărului de voturi pentru ND și GS depinde de mărimea localității. Graficul indică mai degrabă un răspuns negativ: relația este slab pozitivă (un avantaj puțin mai mare pentru ND pe măsură ce crește mărimea localității). Pare că a existat o reacție similară ca direcție și intensitate în cazul majorității localităților. În cazul unor localități intensitatea reacției a fost semnificativ mai mare comparativ cu media. Diferența de intensitate are legătură cu componența etnică a acelor localități (populație de etnie maghiară), nu cu mărimea acestora.

Suplimentar, putem identifica localitățile cu comportament de vot diferit: GS a avut o creștere semnificativ mai mare în Secu (DJ) (nu apare în grafic) și Dognecea (CS); cea mai mare localitate în care GS a avut o creștere mai mare comparativ cu ND este Vicovu de Sus (SV; 5376 votanți la T2; 12764 alegători); ND a avut o creștere semnificativ mai mare în Sânmartin (HR) și Odorheiu Secuiesc; raportat la localitățile mari, ND a avut o creștere mai mare în Târgu-Mureș, Oradea și Cluj-Napoca.





ISBN 978-606-37-2848-8