# `eforensics` Analysis of the Philippines 2022 Presidential and Vice Presidential Elections*

Walter R. Mebane, Jr.†

June 6, 2022

*Thanks to Gary Rozal and Arlene Marie A. Lorica for instigating this analysis and for helpful discussion.

†Professor, Department of Political Science and Department of Statistics, Research Professor, Institute for Political Research, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: wmebane@umich.edu).

**Abstract**

I use clustered precinct data to estimate `eforensics`-frauds to measure the magnitude of malevolent distortions of electors' intentions—frauds—in the 2022 elections for president and vice president in the Philippines. In data from the Transparency Server Bongbong Marcos wins the president election with 31,104,175 votes, and Sara Duterte wins the vice president election with 31,561,948 votes. The `eforensics` statistical model estimates the number of precincts that are `eforensics`-fraudulent and the number of `eforensics`-fraudulent votes counted for either Marcos or Duterte in each precinct. A complication in the election is that an alliance prompted eligible voters to act strategically: strategic behavior can trigger false positive estimates of `eforensics`-frauds. Estimates based on Transparency Server vote count and Project of Precincts eligible voter data ($n = 105649$ precincts with usable data) show that 2,079 precincts and about 416,109 votes for Marcos are produced by malevolent distortions, while 25,144 precincts and approximately 1,710,029 votes for Marcos are produced by a mix of bad acts that manufacture votes and of strategic elector behavior that shifts votes between candidates. For Duterte 2,156 precincts and about 467,816 votes are produced by malevolent distortions, while 11,310 precincts and approximately 796,455 votes for Duterte are produced by a mix of manufactured votes and vote-shifting strategic behavior.

The recent election in the Philippines is of interest for election forensics, as by a wide margin Ferdinand Marcos, Jr., won the presidency in an alliance with Sara Duterte, who won the vice presidency—the son of a former dictator and the daughter of the incumbent populist president (Ratcliffe 2022). The election result is described as "a thumping victory. Marcos Jr secured 31 million votes, according to an unofficial count, thought to be the biggest mandate in decades. Sara Duterte, who ran to be vice-president, also won by a huge margin" (Ratcliffe 2022). The alliance attracted prominent supporters: "The pair were backed up by a host of powerful political names including the former president Gloria Macapagal-Arroyo, who reportedly brokered their alliance, and the family of former president Joseph Estrada. The heavyweights grouped together to oppose the rival candidate, Leni Robredo—a reformist who wanted to pass an anti-dynasty law, and ultimately came second" (Ratcliffe 2022).

The wide margins notwithstanding, suspicions abound about the election. "Political parties are weak in the Philippines, and it is instead influential families who struggle for power, using patronage, vote-buying, intimidation and, at times, violence to keep relatives in office" (Ratcliffe 2022). "Throughout Duterte's term, the Marcos family was building its base, with online disinformation campaigns flooding social media with false stories portraying Marcos Sr's rule as a golden era" (Ratcliffe 2022). Vote buying is reportedly prevalent: "Local media reported that, in the days after the election, stores in some areas sold out of mobile phones, and restaurants were busier than usual, with shoppers openly saying they had received money from politicians running for congress and provincial posts" (Ratcliffe 2022). Some instances of vote buying were officially documented: "The Commission on Elections told local media it had verified more than 100 cases of vote-buying, but did not believe such reports had affected the outcome of the vote" (Ratcliffe 2022).

Given the wide margins in the presidential and vice presidential elections, even a large number of votes produced by malevolent distortions of electors' intentions—or bad acts—might not have changed the outcome of the vote in the sense that the winners might still have the most votes even if fraudulent votes could be removed or correctly compensated. I use precinct data to estimate `eforensics`-frauds (Ferrari, Mebane, McAlister and Wu 2019; Mebane 2022) to measure the magnitude of such votes. Vote count data come from the "Transparency Server" (Commission on Elections 2022*b*; Mendoza 2022).[1] Counts of eligible voters in each precinct come from two sources: first is a variable `NUMBER_VOTERS` in the Transparency file; second is a variable `CLUSTERTOTAL` from Project of Precincts data (Commission on Elections 2022*a*).[2] The `CLUSTERTOTAL` counts are weakly greater than the `NUMBER_VOTERS` counts: 105867 `CLUSTERTOTAL` counts are larger than the corresponding `NUMBER_VOTERS` counts, while for 141 precincts the two kinds of counts are equal. Note that precincts in the data and analysis are actually "clustered precincts" in the terminology of Project of Precincts.

`eforensics` operationalizes the idea that `eforensics`-frauds occur when one candidate gains votes by a combination of manufacturing votes from abstentions and stealing votes from opposing candidates (Mebane 2022). The Bayesian specification of `eforensics` allows posterior means and credible intervals for counts of `eforensics`-fraudulent votes to be determined both for the entire election and for individual precincts. The model requires that some ballot alternative be designated the "leader," which is the alternative that the model allows to benefit from added `eforensics`-fraudulent votes. The candidate with the most votes in each election is this designated leader candidate: `eforensics`-fraudulent votes can add to the votes for Marcos for president and Duterte for vice president.

The most important feature of `eforensics` to keep in mind when considering `eforensics` estimates is that `eforensics` likely responds both to bad acts such as vote-buying, intimidation, violence and disinformation and to strategic elector behavior as

[1]On May 20, 2022, I received an email message from "Partido Liberal (via Google Drive)" containing a URL pointing to a GitHub site that contained Transparency Server data. The key file that contains vote counts, `results.csv`, has a timestamp of May 13 02:50.

[2]Specifically the Project of Precincts data come from `https://comelec.gov.ph/php-tpls-attachments/2022NLE/POP/2022NLEPOP.xlsx`, downloaded on May 9, 2022.

might be entailed in the alliance if voters followed the lead of "heavyweights" and "grouped together" to support the leading candidates and oppose other candidates. A challenge for `eforensics` is to be able to identify which `eforensics`-fraudulent votes reflect malevolent distortions or bad acts and which stem from strategic behavior by electors (eligible voters) (see Mebane 2022). The `eforensics` model distinguishes "incremental frauds" from "extreme frauds": extreme frauds are larger. Analysis like that described in Mebane (2022) suggests that often strategic behavior produces positive incremental frauds estimated `eforensics`-fraudulent vote counts, but usually strategic behavior is not associated with positive extreme frauds estimated counts. So incremental frauds estimates are generally more ambiguous than are extreme frauds estimates.

Table 1 shows the total counts of votes for each candidate and for electors in the raw data from the Transparency Server and from Project of Precincts. Bongbong Marcos has the most votes for president and Sara Duterte has the most votes for vice president. The total for Duterte (31,561,948) somewhat exceeds the total for Marcos (31,104,175). A discrepancy is apparent in the total number of electors between `NUMBER_VOTERS` (55,197,306) and `CLUSTERTOTAL` (66,194,517). The `UNDERVOTE` and `OVERVOTE` counts show the number of ballots that lack valid votes; `eforensics` doesn't do anything with these variables and omits observations that lack valid votes.

Not all precincts have usable data. Project of Precincts includes data for $n = 107785$ (clustered) precincts but the Transparency Server data include only $n = 106008$ precincts. We require that each precinct used to estimate `eforensics` have nonmissing values for the number of electors, the number of votes cast and the number of votes cast for the leader. The number of electors must be weakly greater than the number of votes cast, which must be weakly greater than the number of votes for the leader.

Before considering `eforensics` estimates it is useful to examine plots of the precinct data. Figure 1 shows scatterplots, histograms and empirical densities of precinct data focused on the presidential and vice presidential vote leaders. Plots show turnout (number

Table 1: Philippines 2022 Election Vote and Elector Totals

| Contest | Candidate (Party) or Feature | Count |
|---|---|---|
| President | | |
| | Pacquiao, Manny Pacman (PROMDI) | 3629805 |
| | Montemayor, Jose Jr. (DPP) | 59944 |
| | Mangondato, Faisal (KTPNAN) | 259576 |
| | Domagoso, Isko Moreno (AKSYON) | 1900010 |
| | Marcos, Bongbong (PFP) | 31104175 |
| | Lacson, Ping (PDR) | 882236 |
| | Gonzales, Norberto (PDSP) | 89097 |
| | de Guzman, Leody (PLM) | 92070 |
| | Robredo, Leni (IND) | 14822051 |
| | Abella, Ernie (IND) | 113242 |
| Vice President | | |
| | Atienza, Lito (PROMDI) | 267530 |
| | Serapio, Carlos (KTPNAN) | 89920 |
| | Ong, Doc Willie (AKSYON) | 1851498 |
| | Sotto, Vicente Tito (NPC) | 8183184 |
| | Pangilinan, Kiko (LP) | 9232883 |
| | David, Rizalito (DPP) | 55478 |
| | Lopez, Manny SD (WPP) | 157877 |
| | Bello, Walden (PLM) | 99740 |
| | Duterte, Sara (LAKAS) | 31561948 |
| Eligible Voters and Misvotes | | |
| | `NUMBER_VOTERS` | 55197306 |
| | `CLUSTERTOTAL` | 66194517 |
| | `UNDERVOTE` | 1434939 |
| | `OVERVOTE` | 848815 |

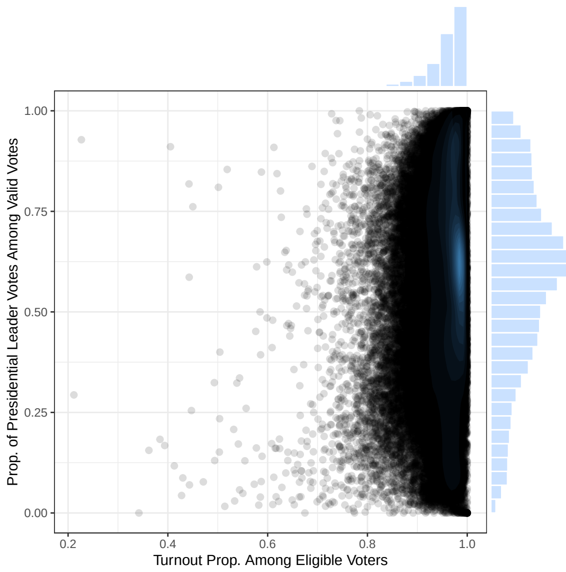
Note: number of voters and vote totals by candidate.

of valid votes divided by number of eligible voters) and leader vote proportions (number voting for leading candidate divided by number of valid votes) in precincts. Figures 1(a,b) use Transparency Server data, while Figures 1(c,d) use Transparency Server data for vote counts but Project of Precincts data for eligible voter counts. Turnout based solely on the data from the Transparency Server is implausibly high: the `NUMBER_VOTERS` variable does not count the number of eligible voters, but rather some subset of the electors. Multimodality is apparent in the vote proportion distributions, and the joint distributions of turnout and vote proportions based on Project of Precincts data are noticeably clumpy.

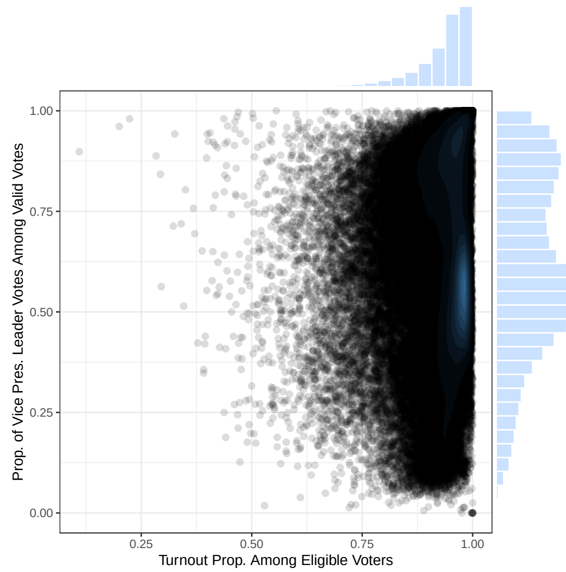
Figure 1: Philippines 2022 Presidential Election Data Plots

Transparency Server vote count and eligible voter data

(a) Marcos by turnout

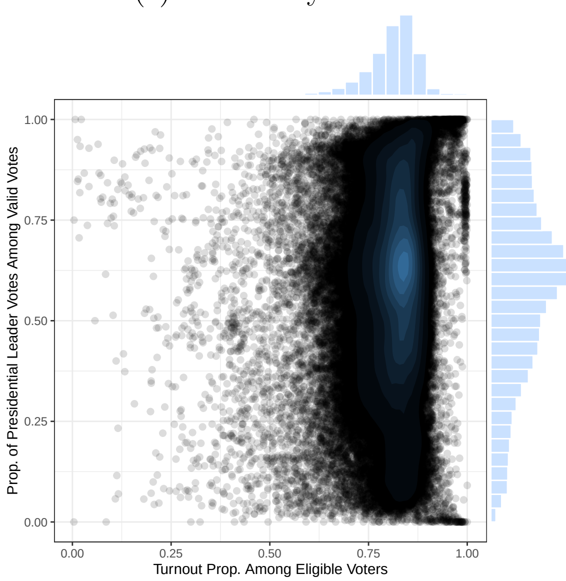



(b) Duterte by turnout

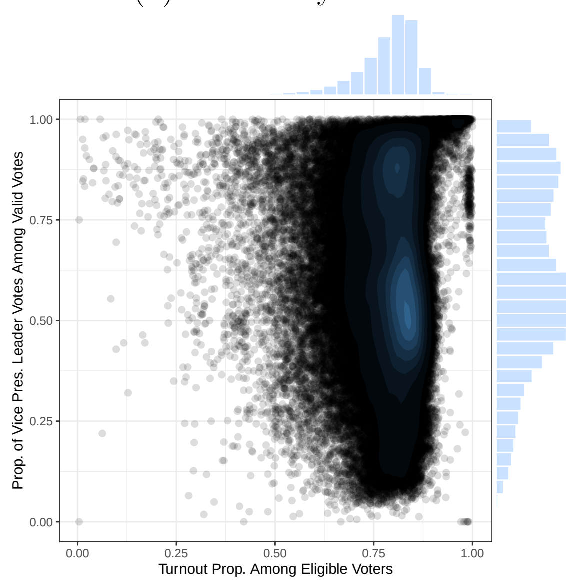



Transparency Server vote count and Project of Precincts eligible voter data

(c) Marcos by turnout




(d) Duterte by turnout



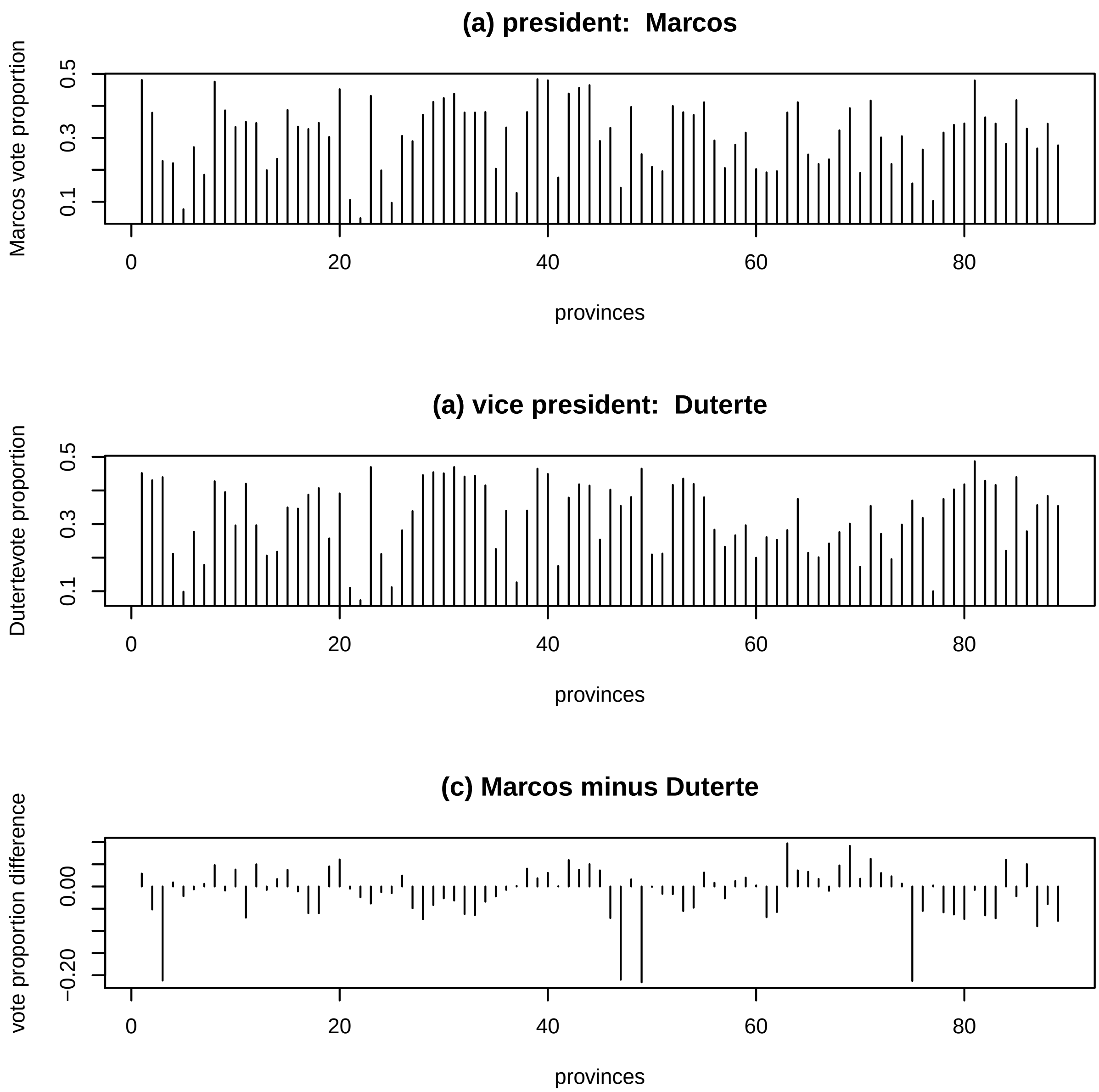

Note: plots show turnout (number voting/number eligible) and vote proportions (number voting for candidate/number voting) for the leading presidential and vice presidential candidates in precincts in the Philippines 2022 elections. Plots show scatterplots with estimated bivariate densities overlaid, with histograms along the axes.

The candidates' support varied regionally: "the candidates combined their respective support bases in the north and south of the country" (Ratcliffe 2022). These regional variations explain some of the multimodal features of Figure 1. Figure 2 shows there is substantial variation in vote proportions for Marcos and for Duterte across provinces.[3] The bimodality in vote proportions that is more pronounced in Figures 1(b,d) than in Figures 1(a,c) traces especially to the four provinces in which Duterte's vote proportion greatly exceeds the proportion for Marcos (Figure 2(c)).

Figure 3 shows plots similar to those in Figure 1 except using turnout and vote proportions expressed as deviations from province means. In Figure 3 the marginal histograms appear closer to Normal than do the histograms in Figure 1, but the scatterplots still exhibit clumpiness. Such clumpiness is a symptom of the kind of dependence among individual electors that informs estimates from the `eforensics` model (Mebane 2022). Malevolent distortions of electors' intentions—frauds—cause such dependence, as can strategic behavior by electors and other election administration failures. Given `eforensics` estimates the principal interpretation question is whether among estimated `eforensics`-frauds malevolent distortions can be discriminated from strategic behavior.

[3]Province names in Figures 2, 4 and 5 corresponding to numbers shown on the $x$-axis of the plots: 1 Abra, 2 Agusan del Norte, 3 Agusan del Sur, 4 Aklan, 5 Albay, 6 Americas, 7 Antique, 8 Apayao, 9 Asia Pacific, 10 Aurora, 11 Basilan, 12 Bataan, 13 Batanes, 14 Batangas, 15 Benguet, 16 Biliran, 17 Bohol, 18 Bukidnon, 19 Bulacan, 20 Cagayan, 21 Camarines Norte, 22 Camarines Sur, 23 Camiguin, 24 Capiz, 25 Catanduanes, 26 Cavite, 27 Cebu, 28 Cotabato, 29 Davao de Oro, 30 Davao del Norte, 31 Davao del Sur, 32 Davao Occidental, 33 Davao Oriental, 34 Dinagat Islands, 35 Eastern Samar, 36 Europe, 37 Guimaras, 38 Ifugao, 39 Ilocos Norte, 40 Ilocos Sur, 41 Iloilo, 42 Isabela, 43 Kalinga, 44 La Union, 45 Laguna, 46 Lanao del Norte, 47 Lanao del Sur, 48 Leyte, 49 Maguindanao, 50 Marinduque, 51 Masbate, 52 Middle East and Africas, 53 Misamis Occidental, 54 Misamis Oriental, 55 Mountain Province, 56 Ncr–Fourth District, 57 Ncr–Manila, 58 Ncr–Second District, 59 Ncr–Third District, 60 Negros Occidental, 61 Negros Oriental, 62 Northern Samar, 63 Nueva Ecija, 64 Nueva Vizcaya, 65 Occidental Mindoro, 66 Oriental Mindoro, 67 Palawan, 68 Pampanga, 69 Pangasinan, 70 Quezon, 71 Quirino, 72 Rizal, 73 Romblon, 74 Samar, 75 Sarangani, 76 Siquijor, 77 Sorsogon, 78 South Cotabato, 79 Southern Leyte, 80 Sultan Kudarat, 81 Sulu, 82 Surigao del Norte, 83 Surigao del Sur, 84 Tarlac, 85 Tawi-Tawi, 86 Zambales, 87 Zamboanga del Norte, 88 Zamboanga del Sur, 89 Zamboanga Sibugay.

Figure 2: Philippines 2022 Vote Proportions by Province

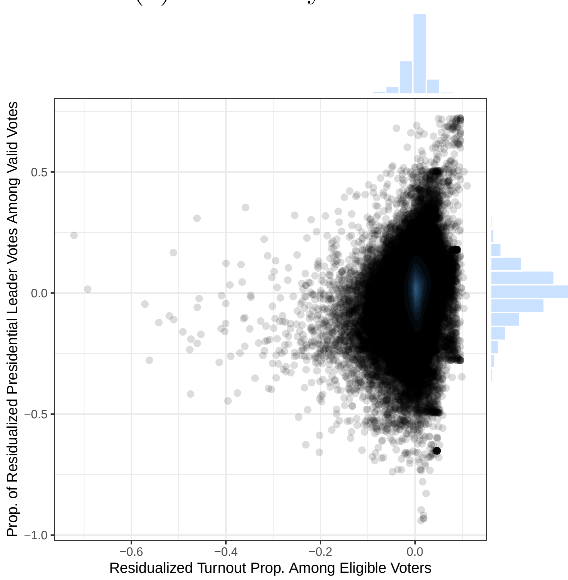



Note: vote proportions by province for Marcos and Duterte.

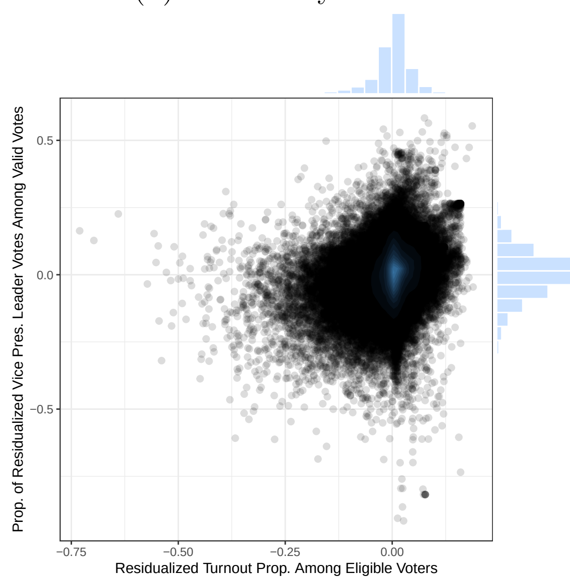
Figure 3: Residualized Philippines 2022 Presidential Election Data Plots

Transparency Server vote count and eligible voter data

(a) Marcos by turnout

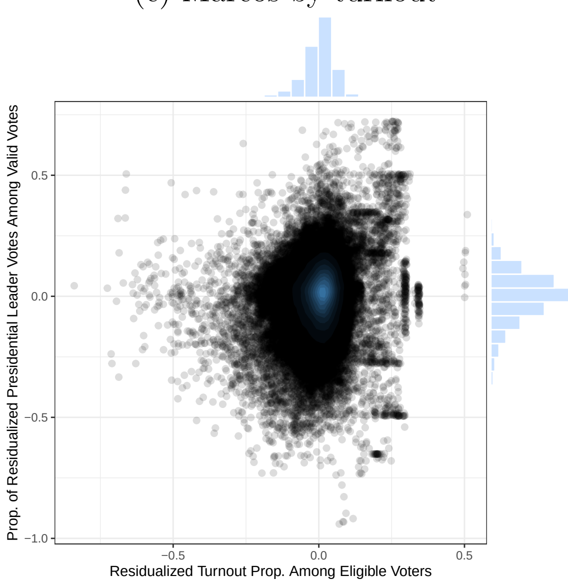



(b) Duterte by turnout

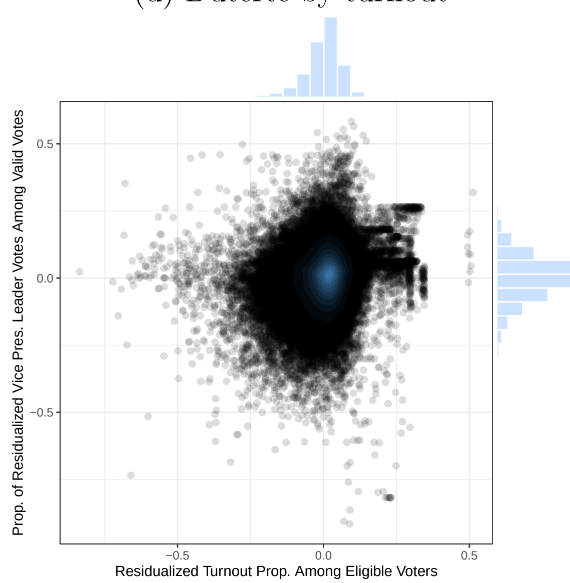



Transparency Server vote count and Project of Precincts eligible voter data

(c) Marcos by turnout




(d) Duterte by turnout




Note: plots show turnout (number voting/number eligible) and vote proportions (number voting for candidate/number voting) for the leading presidential and vice presidential candidates in precincts in the Philippines 2022 elections. Plots show scatterplots with estimated bivariate densities overlaid, with histograms along the axes.

`eforensics` estimates for president based on Transparency Server vote count and Project of Precincts eligible voter data are in Table 2.[4] The `eforensics` model specification includes province fixed effects as covariates for the turnout and vote choice linear predictors.[5] The table shows `eforensics` parameters, the total number of precincts classified as `eforensics`-fraudulent and the number of `eforensics`-fraudulent votes. For the parameters the table reports posterior means and 95% highest-posterior density (HPD) intervals. Among the precincts classified as `eforensics`-fraudulent, the numbers with incremental versus extreme fraud are reported. The number of manufactured votes ($F_t$) is shown separately from the total number of `eforensics`-fraudulent votes ($F_w$); the number of stolen votes is $F_w - F_t$. For $F_t$ and $F_w$ the tables report posterior means and 99.5% credible intervals.

Table 2: Philippines 2022 President Election `eforensics` Estimates, Province Fixed Effects

| Type | Parameter | Covariate | Mean | lo$^a$ | up$^b$ |
|---|---|---|---|---|---|
| mixture probabilities | $\pi_1$ | No Fraud | .678 | .579 | .886 |
| | $\pi_2$ | Incremental Fraud | .301 | .107 | .398 |
| | $\pi_3$ | Extreme Fraud | .0204 | .00761 | .0278 |
| incremental frauds | $\rho_{M0}$ | (Intercept) | $-.194$ | $-.645$ | .544 |
| | $\rho_{S0}$ | (Intercept) | $-.424$ | $-.913$ | $-.146$ |
| extreme frauds | $\delta_{M0}$ | (Intercept) | $-2.53$ | $-4.51$ | $-.302$ |
| | $\delta_{S0}$ | (Intercept) | $-.189$ | $-.649$ | .189 |

units fraudulent: 27223 fraudulent (25144 incremental, 2079 extreme), 78426 not fraudulent
manufactured votes $F_t = 1078593.5\ [563278.7, 2299447.9]^c$
total fraudulent votes $F_w = 2126137.5\ [806628.7, 3329843.1]^c$

Note: `eforensics` model parameter estimates (posterior means and credible intervals). Province fixed effects not shown. $n = 105649$ precinct units. $^a$ 95% HPD lower bound. $^b$ 95% HPD upper bound. $^c$ posterior mean [99.5% credible interval]. Transparency Server vote count and Project of Precincts eligible voter data.

`eforensics` parameters support interpretations regarding whether the data feature malevolent distortions in votes (i.e., bona fide frauds), strategic behavior by electors or failures of election administration (Mebane 2022). The sizes of $\pi_2$ and $\pi_3$ or of $F_t$ and $F_w$

[4] Each Monte Carlo Markov Chain (MCMC) estimation uses four chains.

[5] These covariates are included in in $x^\tau$ and $x^\nu$ of Mebane (2022, equation (2a–2b)).

are not good indicators in themselves for whether `eforensics` estimates are responding to malevolent distortions, to strategic behavior, to administrative failures or what. But the parameters are the guides to interpreting `eforensics`-frauds: incremental frauds intercept parameters $\rho_{M0}$ and $\rho_{S0}$ are often negative when only strategic behavior is occurring; multimodality in the posterior distribution of mixture probability $\pi_2$ may indicate there are lost votes (Mebane 2022).

The estimates in Table 2 show many precincts have `eforensics`-frauds: 27,223 of 105,649 precincts in the analysis are classified as `eforensics`-fraudulent. The posterior mean estimate for the total number of `eforensics`-fraudulent votes for Marcos is 2126137.5, which is smaller than the gap of 16,282,124 between Marcos and the second-place candidate.

With these results and with the results to be reported for vice president, a key question is whether the estimated `eforensics`-frauds reflect malevolent distortions, strategic elector behavior, administrative failures or what. The mixture probability estimates exhibit multimodality that may indicate there are lost votes (Mebane 2022): the wide credible interval for $\pi_2$ in Table 2 signals multimodality that is apparent among the four Markov Chains used for estimation;[6] if there are lost votes, whether they result from election administration problems or from voter suppression cannot be determined merely from the available count data.

To distinguish between malevolent distortions and strategic elector behavior I focus on the distinction between incremental and extreme `eforensics`-frauds. As I mentioned previously, often strategic behavior produces positive estimated incremental frauds counts, but usually strategic behavior is not associated with positive estimated extreme frauds counts. So incremental frauds estimates are generally more ambiguous than are extreme frauds estimates. With reference to Table 2 this suggests that the estimated extreme frauds can reliably be attributed to malevolent distortions, while the estimated incremental

[6]Separately across the four chains, posterior means and 95% HPD intervals are $\pi_2 = .394$ [.389, .399], $\pi_2 = .110$ [.106, .115], $\pi_2 = .307$ [.304, .310], $\pi_2 = .393$ [.388, .399].

frauds might be triggered partly by strategic behaviors (Mebane 2022) such as wasted-vote behavior and behavior to effect the alliance in which elites urged electors to "group together" to support the leading candidates and oppose other candidates. In other elections incremental frauds intercept parameters $\rho_{M0}$ and $\rho_{S0}$ are often negative when only strategic behavior is occurring: in Table 2 $\rho_{S0} < 0$ suggests that incremental fraud vote stealing in the estimates probably relates at least partially to strategic behavior, while the indeterminate sign of $\rho_{M0}$ suggests that incremental fraud manufactured votes trace to malevolent distortions.

It is notable that the number of precincts that have incremental `eforensics`-frauds exceeds the number of precincts that have extreme `eforensics`-frauds by a factor of more than ten: 25,144 precincts have incremental frauds while 2,079 precincts have extreme frauds. The average number of `eforensics`-fraudulent votes in `eforensics`-fraudulent precincts is 118.0 (60.9 manufactured) given incremental fraud and 239.4 (110.8 manufactured) given extreme fraud. A posterior mean of 1,710,029 of the `eforensics`-fraudulent votes come from incremental frauds and 416,109 come from extreme frauds.[7] So a cautious intepretation of the estimates says that 2,079 precincts and about 416,109 votes for Marcos are produced by bad acts—more by stealing votes from other candidates than by manufacturing votes from nonvoters—while many of the 25,144 precincts and approximately 1,710,029 votes for Marcos that have incremental `eforensics`-frauds are produced a mix of bad acts that manufacture votes and of strategic elector behavior that shifts votes between candidates.

It is reasonable to say that many of those 1,710,029 incrementally `eforensics`-fraudulent votes are prompted by efforts to enact the alliance. The evidence for this is a contrast with 2016 president `eforensics` estimates produced using a specification similar to that used for Table 2 except omitting province fixed effects.[8] In the

[7]Posterior means and 99% credible intervals for total number of votes from incremental frauds and extreme frauds are, respectively, 1710028.8 $[550294.3, 2815604.7]$ and 416108.6 $[256334.4, 514709.1]$.

[8]Note that the data used to estimate the `eforensics` for 2016 include only 89,823 precincts and a total of only 52,979,970 registered voters, 41,068,434 votes cast for president and 15,776,653 votes for the leading

2016 president election 88,390 precincts are not `eforensics`-fraudulent and 1,432 precincts are `eforensics`-fraudulent, with 962 precincts having incremental frauds and 470 having extreme frauds. The posterior mean for the total number of `eforensics`-fraudulent votes for Rodrigo Duterte is 225,793. Rodrigo Duterte received 39.02% of the votes while the next three finishers received 23.45%, 21.39% and 12.73%. Certainly in the 2016 election there was no alliance: Bongbong Marcos finished second running for vice president with 34.37% of the votes, behind Leni Robredo who had 35.11% (the third- and fourth-place candidate had 14.38% and 12.01%). While the 2022 presidential election has 4.4 times more precincts with extreme frauds than does the 2016 election, it has 26.1 times more precincts with incremental frauds.[9] Most likely the difference is the 2022 alliance.

Notice that the 2022 Philippines `eforensics` estimates resemble those reported in Mebane (2022) for the Turkey 2011 legislative election: in that election 70,155 polling stations have no fraud while 129,400 have `eforensics`-frauds; 123,549 polling stations have incremental frauds and 5,851 have extreme frauds; a posterior mean of 3,639,965 votes received by leading legislative candidates are estimated to be `eforensics`-fraudulent; these outcomes likely reflect that many candidates that year ran as independents (Kesgin 2012), which reduced party signals and prompted more strategic coordination among electors; the increased role for strategic elector behavior sharply increased the prevalence of incremental frauds.

Table 3 reports estimates for vice president based on Transparency Server vote count and Project of Precincts eligible voter data, using an `eforensics` specification similar to that used for the president election as reported in Table 2—the `eforensics` model specification includes province fixed effects as covariates for the turnout and vote choice linear predictors. For the vice president election 13,466 precincts are `eforensics`-fraudulent (11,310 incremental, 2,156 extreme). The posterior mean of the

presidential candidate Rodrigo Duterte.

[9]Taking the larger number of precincts in the analysis in 2022 into account, 2002 has 3.9 times more precincts with extreme fraud than does 2016, and 22.2 times more precincts with incremental fraud.

total number of `eforensics`-fraudulent votes for Marcos is larger than that for Duterte—2,126,138 versus 1,264,271—but the 99.5% credible intervals for these totals overlap, so it is reasonable to say the values are similar. In the estimates for the vice president election $\pi_2$ has a wide HPD interval that traces to a multimodal posterior,[10] which may suggest votes were lost in the vice president election. The intercept coefficients in the frauds linear predictors[11] are similar to those in the president election. in Table 3 $\rho_{S0} < 0$ suggests that incremental fraud vote stealing in the estimates probably relates at least partially to strategic behavior, while the indeterminate sign of $\rho_{M0}$ suggests that incremental fraud manufactured votes trace to malevolent distortions.

Table 3: Philippines 2022 Vice President Election `eforensics` Estimates, Province Fixed Effects

| Type | Parameter | Covariate | Mean | lo[a] | up[b] |
|---|---|---|---|---|---|
| mixture probabilities | $\pi_1$ | No Fraud | .766 | .726 | .863 |
| | $\pi_2$ | Incremental Fraud | .213 | .126 | .251 |
| | $\pi_3$ | Extreme Fraud | .0205 | .0106 | .0258 |
| incremental frauds | $\rho_{M0}$ | (Intercept) | $-.164$ | $-.682$ | .510 |
| | $\rho_{S0}$ | (Intercept) | $-.584$ | $-.934$ | $-.187$ |
| extreme frauds | $\delta_{M0}$ | (Intercept) | $-1.31$ | $-2.38$ | $-.334$ |
| | $\delta_{S0}$ | (Intercept) | $-.207$ | $-.531$ | .163 |

units fraudulent: 13466 fraudulent (11310 incremental, 2156 extreme), 92183 not fraudulent
manufactured votes $F_t = 858873.4$ $[383837.6, 1794782.0]^c$
total fraudulent votes $F_w = 1264270.9$ $[743204.0, 2387093.0]^c$

Note: `eforensics` model parameter estimates (posterior means and credible intervals). Province fixed effects not shown. $n = 105649$ precinct units. [a] 95% HPD lower bound. [b] 95% HPD upper bound. [c] posterior mean [99.5% credible interval]. Transparency Server vote count and Project of Precincts eligible voter data.

The `eforensics` estimates reported in Table 3 suggest that both malevolent distortions and strategic elector behavior affected the elections' results. A posterior mean of 796,455 of the `eforensics`-fraudulent votes come from incremental frauds and 467,816 come from

[10]Separately across the four chains, posterior means and 95% HPD intervals are $\pi_2 = .237$ [.229, .246], $\pi_2 = .241$ [.238, .244], $\pi_2 = .128$ [.126, .131], $\pi_2 = .247$ [.233, .253].

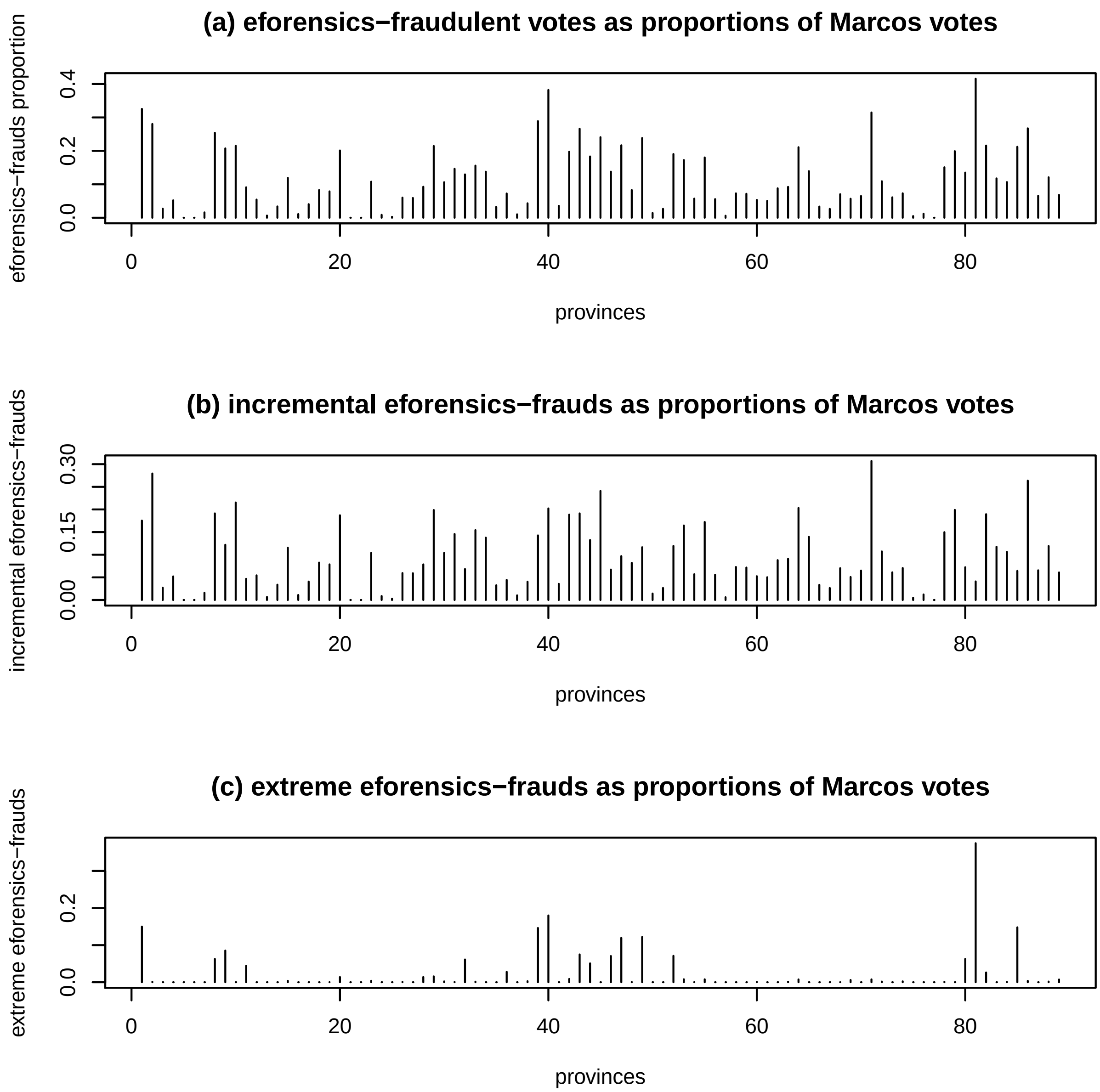
[11]See Mebane (2022, equation (2c–2d)).

extreme frauds.[12] A cautious intepretation of the estimates says that 2,156 precincts and about 467,816 votes for Duterte are produced by bad acts, while the 11,310 precincts and approximately 796,455 votes for Duterte that have incremental `eforensics`-frauds are produced by a mix of bad acts that manufacture votes and of strategic elector behavior that shifts votes between candidates.

Figures 4 and 5 show precinct `eforensics`-fraudulent vote estimates, aggregated to province totals, as proportions of the total vote for the leading candidates in each precinct. The province sums of the `eforensics`-fraudulent vote estimates are sums of the posterior mean for each precinct. Combining all `eforensics`-fraudulent vote proportions for each province the proportions are somewhat similar between Marcos and Duterte: the product-moment correlation between the proportions shown in Figures 4(a) and 5(a) is $r = .64$. Considering separately incremental and extreme frauds the similarities are lower for incremental frauds and greater for extreme frauds: respectively, $r = .43$ and $r = .86$. Several provinces have high proportions of extreme frauds for both Marcos and Duterte: these include provinces 1 (Abra), 8 (Apayao), 39 (Ilocos Norte), 49 (Maguindanao), 52 (Middle East and Africas), 80 (Sultan Kudarat), 81 (Sulu) and 85 (Tawi-Tawi). One province where Marcos has notably less support than does Duterte (Figure 2(c)) also has substantial extreme frauds proportions for Duterte but not for Marcos: province 47 (Lanao del Sur). Extreme `eforensics`-frauds may account for much of the difference between Duterte and Marcos in that province.

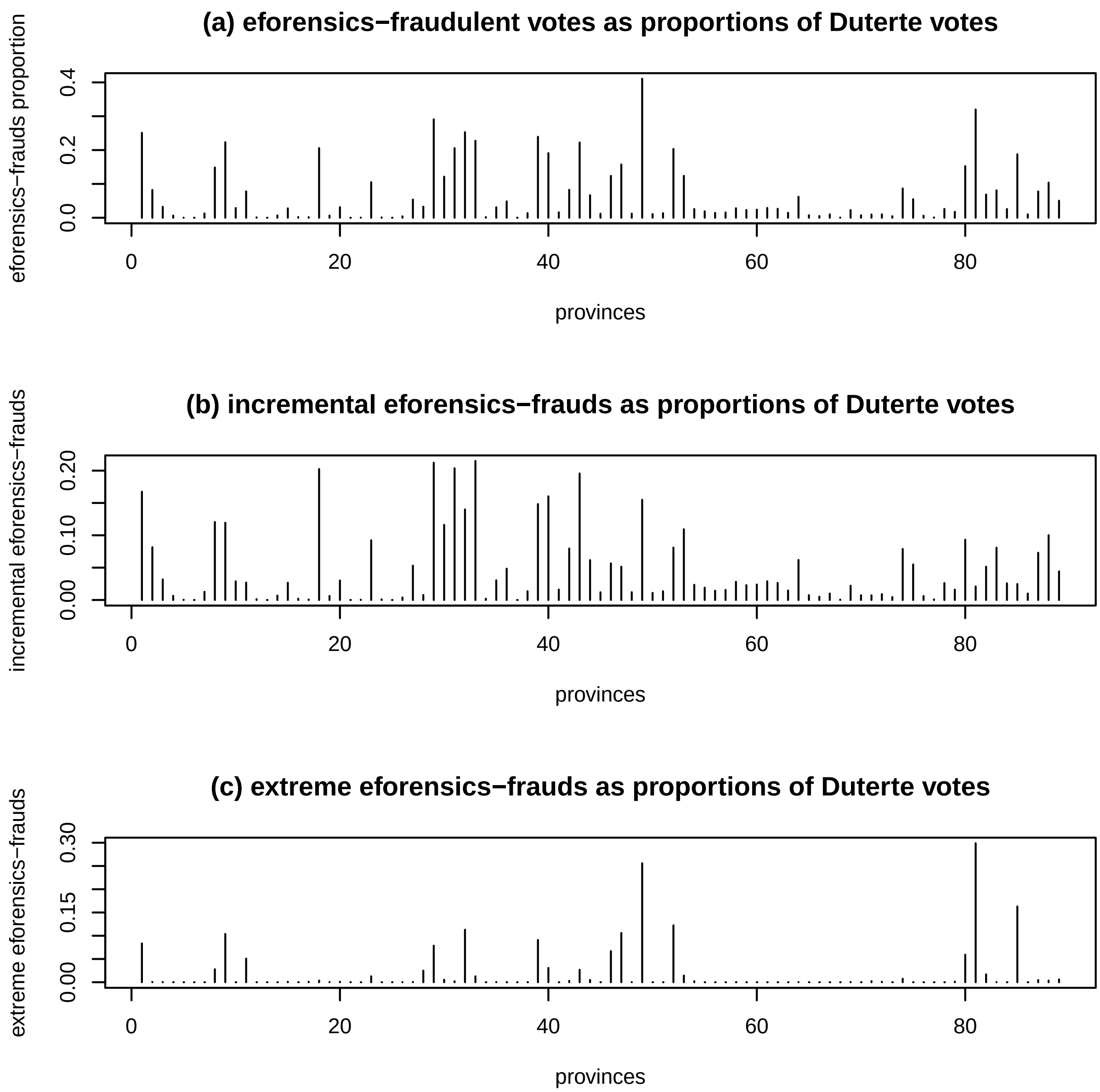
[12] Posterior means and 99% credible intervals for total number of votes from incremental frauds and extreme frauds are, respectively, 796455.0 $[310522.3, 1745006.9]$ and 467815.8 $[338591.8, 643738.9]$.

Figure 4: Philippines 2022 President Election `eforensics`-frauds Proportions by Province




Note: plots show province totals of precinct `eforensics`-fraudulent vote posterior mean estimates divided by the province sum of the number voting for the leading candidate. (a) all `eforensics`-frauds. (b) incremental `eforensics`-frauds. (c) extreme `eforensics`-frauds. Estimates come from a model specification that includes province fixed effects for turnout and vote choice. Transparency Server data.

Figure 5: Philippines 2022 Vice Pres. Election `eforensics`-frauds Proportions by Province

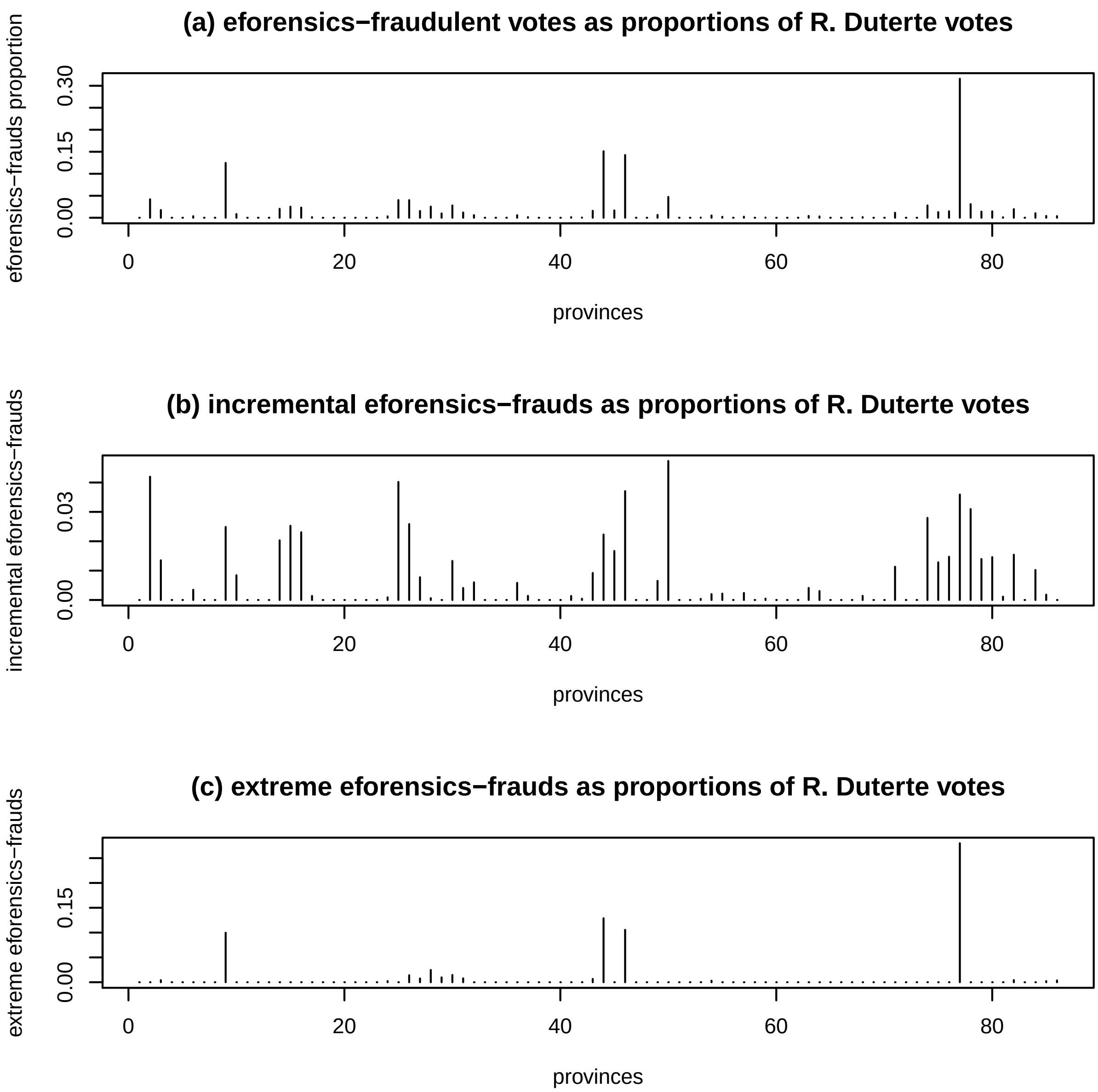



Note: plots show province totals of precinct `eforensics`-fraudulent vote posterior mean estimates divided by the province sum of the number voting for the leading candidate. (a) all `eforensics`-frauds. (b) incremental `eforensics`-frauds. (c) extreme `eforensics`-frauds. Estimates come from a model specification that includes province fixed effects for turnout and vote choice. Transparency Server data.

To faciliatate comparisons with the 2016 president election, Figure 6 shows precinct `eforensics`-fraudulent vote estimates, aggregated to province totals, as proportions of the total vote for the leading 2016 presidential candidate in each precinct.[13] High proportions of extreme frauds are apparent in provinces 9 (Basilan), 44 (Lanao del Sur), 46 (Maguindanao) and 77 (Sulu).

[13] In Figure 6 numbers on the $x$-axis of the plot correspond to provinces as follows: 1, Abra; 2, Agusan del Norte; 3, Agusan del Sur; 4, Aklan; 5, Albay; 6, Antique; 7, Apayao; 8, Aurora; 9, Basilan; 10, Bataan; 11, Batanes; 12, Batangas; 13, Benguet; 14, Biliran; 15, Bohol; 16, Bukidnon; 17, Bulacan; 18, Cagayan; 19, Camarines Norte; 20, Camarines Sur; 21, Camiguin; 22, Capiz; 23, Catanduanes; 24, Cavite; 25, Cebu; 26, Compostela Valley; 27, Cotabato (North Cot.); 28, Davao (Davao del Norte); 29, Davao del Sur; 30, Davao Occidental; 31, Davao Oriental; 32, Dinagat Islands; 33, Eastern Samar; 34, Guimaras; 35, Ifugao; 36, Ilocos Norte; 37, Ilocos Sur; 38, Iloilo; 39, Isabela; 40, Kalinga; 41, La Union; 42, Laguna; 43, Lanao del Norte; 44, Lanao del Sur; 45, Leyte; 46, Maguindanao; 47, Marinduque; 48, Masbate; 49, Misamis Occidental; 50, Misamis Oriental; 51, Mountain Province; 52, National Capital Region–Fourth District; 53, National Capital Region–Manila; 54, National Capital Region–Second District; 55, National Capital Region–Third District; 56, Negros Occidental; 57, Negros Oriental; 58, Northern Samar; 59, Nueva Ecija; 60, Nueva Vizcaya; 61, Occidental Mindoro; 62, Oriental Mindoro; 63, Palawan; 64, Pampanga; 65, Pangasinan; 66, Quezon; 67, Quirino; 68, Rizal; 69, Romblon; 70, Samar (Western Samar); 71, Sarangani; 72, Siquijor; 73, Sorsogon; 74, South Cotabato; 75, Southern Leyte; 76, Sultan Kudarat; 77, Sulu; 78, Surigao del Norte; 79, Surigao del Sur; 80, Taguig–Pateros; 81, Tarlac; 82, Tawi-Tawi; 83, Zambales; 84, Zamboanga del Norte; 85, Zamboanga del Sur; 86, Zamboanga Sibugay.

Figure 6: Philippines 2016 Pres. Election `eforensics`-frauds Proportions by Province




Note: plots show province totals of precinct `eforensics`-fraudulent vote posterior mean estimates divided by the province sum of the number voting for the leading candidate. (a) all `eforensics`-frauds. (b) incremental `eforensics`-frauds. (c) extreme `eforensics`-frauds. Estimates come from a model specification that includes province fixed effects for turnout and vote choice. Transparency Server data.

# References


Commission on Elections, Republic of the Philippines. 2022*a*. "2022 National and Local Elections Project of Precincts." Apr 11, 2022, URL: `https://comelec.gov.ph/?r=2022NLE/ProjectOfPrecincts`.

Commission on Elections, Republic of the Philippines. 2022*b*. "Guidelines on the Issuance of Access to the Election Results Data from the Media Server in the University of Sto. Thomas Relative to the Conduct of the May 09, 2022 National and Local Elections." URL: `https://comelec.gov.ph/php-tpls-attachments/2022NLE/Resolutions/com_res_10781.pdf`.

Ferrari, Diogo, Walter Mebane, Kevin McAlister and Patrick Wu. 2019. *Election Forensics: Positive Empirical Models of Election Fraud.* R package version 0.0.4 (Supported by NSF grant SES 1523355). Initial version: August 27, 2019. URL: `https://github.com/UMeforensics/eforensics_public`.

Kesgin, Barş. 2012. "The 2011 Parliamentary Election in Turkey." *Electoral Studies* 31(4):867–870.

Mebane, Jr., Walter R. 2022. "Measuring Election Frauds." Working paper, URL: `http://www.umich.edu/~wmebane/measfrauds.pdf`.

Mendoza, John Eric. 2022. "LIST: 20 Media Outlets Granted Server Access for 2022 Poll Results by COMELEC." *Inquirer.NET* . May 3, 2022, URL: `https://newsinfo.inquirer.net/1591823/list-20-media-outlets-granted-server-access-for-2022-poll-results-by-comelec`.

Ratcliffe, Rebecca. 2022. "Philippines election result is a win for dynasty politics." *The Guardian* . May 26, 2022, URL: `https://www.theguardian.com/world/2022/may/26/philippines-election-result-is-a-win-for-dynasty-politics`.