On the validity of vote counts published by the Venezuelan opposition

Dorothy Kronick*
August 22, 2024[†]

Draft.

Abstract

I investigate data published by the campaign of Edmundo González Urrutia, principal opposition candidate in Venezuela's July 28 presidential election. The campaign claims that these data reflect actual votes cast on election day, while the Venezuelan government claims that the data are fake and that incumbent Nicolás Maduro won the election. I show that the double paper trail generated by Venezuela's electronic voting system—a ballot receipt for each voter and a tally sheet for each voting machine—all but rules out fraud or fabrication on the part of the campaign, just as it has all but ruled out government fraud or fabrication in many prior elections. I conclude that the campaign data almost certainly reflect actual votes cast.

^{*}Assistant Professor, Goldman School of Public Policy, U.C. Berkeley.

[†]For comments, I thank Guy Grossman, Nicolás Idrobo, Jennifer McCoy, Nicolás Idrobo, Dalson Figueiredo, Walter Mebane, Francisco Monaldi, Fernando Neisser, Raphael Nishimura, Francisco Rodríguez, Uri Simonsohn, and Harold Trinkunas.

The U.S. government and the mainstream U.S. media agree that Edmundo González Urrutia defeated incumbent Nicolás Maduro in Venezuela's recent presidential election, despite Maduro's claim to have won. But the governments of Brazil, Colombia, and Mexico—all of which could play a key role in liaising with Maduro—have yet to take a stand, and major left media outlets report that the González campaign is attempting a coup (Democracy Now!, 2024), in cahoots with "the international fascist far right and the CIA" (Counterpunch, 2024).

This skepticism is understandable. The González campaign published data indicating that he won 7.3 million votes within the 83% of voting booths for which the campaign was able to collect tally sheets; if so, he likely won nearly 8.5 million votes in the electorate overall.² That's half a million more votes than Hugo Chávez obtained in the 2012 presidential election—before one-fifth of Venezuelans emigrated, and at a time when the government had the resources to ramp up pre-election spending (Rodríguez, 2024b). The González campaign data also suggest that Maduro won approximately 4.1 million votes nationwide, which is not many more than Maduro's political party got in the 2021 gubernatorial elections—even though the Venezuelan economy has improved since 2021, and even though presidential elections typically mobilize more voters than regional elections. And while public opinion polls did predict a González landslide, those same polls had overestimated opposition vote share in recent elections (Rodríguez, 2024c). For these reasons, a person need not distrust María Corina Machado herself (Ellner, 2024) in order to harbor doubt about the claims of the González campaign: if true, these electoral results are extraordinary.

I evaluate competing claims of victory from Maduro and from the González campaign, finding evidence in support of the campaign's claims. Venezuela's electronic voting system, which Jimmy Carter called "the best in the world," prints paper ballot receipts that, in and of themselves, provide proof of what actually happened on election day. As a result, it is not necessary to trust the voting-machine vendor, or the campaign, or for that matter a single Venezuelan opposition leader, in order to see that Venezuelan voters almost certainly did elect González on July 28.

 $^{^1\}mathrm{The}$ New York Times, 8/1/2024, 8/6/2024; The Washington Post, 8/4/2024; The Wall Street Journal, 8/2/2024

²According to estimates that I discuss later in this research note.

I first describe Venezuela's electronic voting system (Section 1) and then review how it has worked to establish correct vote counts in Venezuelan elections in the past (Section 2). Turning to this year's presidential election, I first show that, if we take data published by the González campaign at face value, that data unambiguously shows that González won with a large margin (Section 3). I then walk through the reasons to credit the campaign data: it derives from Venezuela's electronic voting machines, and the integrity of the electronic voting machines is guaranteed by the record of counted-by-hand paper ballot receipts (Section 4). I specify what we would need to believe, given these ballot receipts, in order to assert that the campaign data are fraudulent or fabricated.

Finally, I consider the potential role of the paper ballot receipts in possible upcoming discussions (Section 5). As of this writing, three weeks after the election, it seems that the Venezuelan electoral council may publish data that appears to provide evidence of Maduro's win; in that case, the record of counted-by-hand paper ballots should help establish the truth of what happened on election day.

1 Venezuela's electronic voting system

Venezuela's electronic voting system begins with fingerprint-based voter identification. Each voter shows her national ID card (expired cards are explicitly permitted), types her national ID number ($c\acute{e}dula$) into a keypad, and places her right thumb on a fingerprint scanner (CNE, 2024). If her scanned thumbprint matches her thumbprint in Venezuela's national fingerprint registry, the match activates her electronic voting machine and the voter proceeds to her voting booth (mesa).

If her thumbprint does not match, poll workers attempt to confirm her identity with other fingerprints (index finger, for example); if this fails, the president of the voting booth can manually override the block in order to activate the voting machine and allow the voter to proceed. But three such manual overrides in the same voting booth trigger an additional block that requires the voting-booth president to enter a special override code—and the entry of three such override codes (i.e., nine votes without fingerprint matches) blocks the voting machine until the voting-booth president can reach the electoral council technical support team for a higher-level override. The

electoral council records the number of no-fingerprint-match votes at each voting machine and later distributes those counts to the political parties.

This system guards against partisan poll workers casting ballots in place of voters who choose to abstain (a form of electronic ballot-box stuffing) and against one voter impersonating others.³ By conditioning voting-machine-activation on a valid finger-print match (or manual overrides of escalating difficulty), the Venezuelan system complicates any attempt to cast a ballot that belongs to someone else. There *are* documented incidents of voter impersonation—as when one Venezuelan arrived at different voting booths with different fake ID cards and obtained manual overrides in order to cast multiple ballots (La Patilla, 2013)—but they are few and far between.

Poll workers in Venezuela are randomly selected from among each booth's registered voters, and selected voters are required to serve (as in Bolivia, and as in jury duty in the United States). Each voting booth has a president, a secretary, and two additional poll workers. Each political candidate may also designate a witness to observe proceedings at each voting booth.

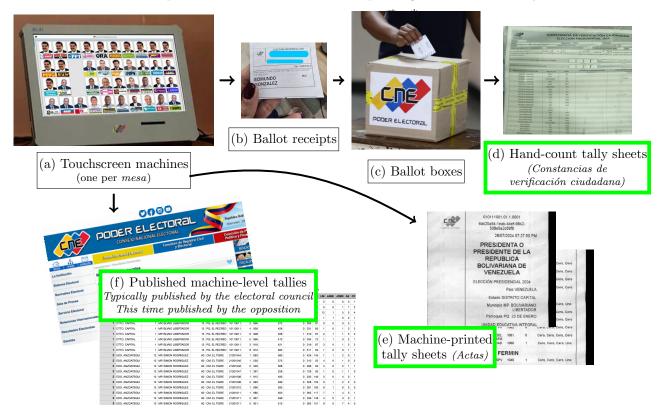
Voters who showed up to the polls in Venezuela on July 28, 2024, encountered a boxy white touchscreen in the style of a mid-2000s iMac (Figure 1a). After selecting their preferred candidate—party pairing, the screen prompted them to confirm their choice or to go back and correct it. The combination of (i) a screen with candidate photos and (ii) the follow-up question likely explains the very low number of invalid votes in Venezuelan elections; in Brazil, for example, the introduction of these two features dramatically reduced the proportion of ballots that were spoiled (Fujiwara, 2015). It is easier to fill in the wrong bubble on a scantron or to leave a chad hanging than to accidentally finger the wrong candidate's face on a screen.

After a voter confirms her choice, the machine ejects a ballot slip (Figure 1b). The voter checks the slip and deposits it in a ballot box. After the last vote is cast, each voting machine prints a tally sheet (acta): a long receipt listing the number of votes for each candidate—party pairing, along with information identifying the voting booth and the poll workers (Figure 1e). Only after the tally sheet is printed does

³The Venezuelan electoral registry itself guards against one person voting in multiple locations using her own identity, as very occasionally does happen in the United States, where there is no national voter roll (Goel et al., 2020).

Figure 1: The paper trail in Venezuela's electronic voting system

Venezuela's electronic voting system produces a double paper trail: one based on individual voters' paper ballot receipts, a second based on tally sheets that each machine prints at the close of voting (actas). When hand counts of paper ballot receipts, machine-printed tally sheets, and voting-machine-level tallies published online all match, this triple congruence rules out tally fraud.



the machine transmit its results to Venezuela's electoral council. Another virtue of Venezuela's electronic voting system is that the tally sheets include a QR code at the bottom that spits out a spreadsheet-able text version of the tallies: no need for data entry or glitchy OCR.

Venezuela allows each candidate to install one witness at each polling place. One of the primary responsibilities of these witnesses is to observe the count of paper ballots that is supposed to take place after the close of voting. Electoral council regulations require poll workers to randomly select approximately half of ballot boxes,⁴ open

⁴ Specifically, a manual published by the electoral council for the 2024 presidential election

them, hand-count the paper ballot receipts, record the tallies on a physical paper form (Figure 1d), and check that the hand-counted totals match those on the machine-printed tally sheet. This step is crucial. Without it, there would be nothing to stop the government (or a hacker, or the opposition, or anyone with access) from rigging the machines to produce any desired result. The paper ballot receipts prevent this. If the machines were rigged, either voters would notice that their paper ballot slips didn't match their votes—wait, I didn't vote for so-and-so—or the hand-counted paper ballot receipts wouldn't match the machine-printed tally sheets.

Venezuelan law also requires the government to publish electoral returns at the level of the voting machine (Figure 1f).⁵ In the twenty years since the electronic voting system was installed, the electoral council has complied in every election save three: the election of a constituyente in 2017, a referendum in 2023, and the presidential election on July 28 of this year, for which the electoral council has (as of this writing) published no data at the level of the voting machine. In all other previous elections, voting-machine-level data published on the website of the electoral council allowed candidates and the public to check whether the totals matched those of the paper trail. If, for example, the website reported that Chávez received 500 votes and the opposition candidate 300 votes in a given voting machine in a given presidential election, people could check that these totals matched both the printed paper tally sheet from that machine and the hand-count of paper ballot slips (if that machine had been selected for hand counting).

A match across all three of these counts, or triple congruence (Toro, 2013), essentially guarantees that votes were counted correctly on election day. In principle, this system provides strong protection against undetected tally fraud. The next section considers how the system has worked in practice.

instructs poll workers to open and hand-count one box in polling places with 1–4 voting machines, two boxes in places with 5–8 voting machines, three boxes in places with 9–11 voting machines, and four boxes in places with more than 11 voting machines. Given the distribution of number of voting machines per polling place, this amounts to 55.4% of boxes. The proportion was similar in prior elections.

⁵Specifically, Article 150 of the *Ley Orgánica de Procesos Electorales* requires the electoral council to publish voting-machine-level data, Article 155 requires the electoral council to do so within 30 days, and Article 148 requires the electoral council to present the total (overall) vote count within 48 hours of the election.

2 Venezuela's electronic voting system in previous elections

One might wonder why a president like Hugo Chávez, who sought to concentrate power in the presidency and perpetuate himself in office (Brewer-Carías, 2010; Corrales and Penfold, 2011), would install an electronic voting system like the one just described. He could have chosen instead riggable voting machines that leave no paper trail, giving the government room to stealthily alter tallies. But Chávez selected Venezuela's electronic voting system at a time when he enjoyed a majority, thanks to his charisma (Hawkins, 2010), lavish consumption spending (Rodríguez, 2008; Gulotty and Kronick, 2022), voter intimidation and other forms of pre-election manipulation (Corrales, 2020), and errors in opposition strategy (Gamboa, 2022). As a result, Chávez didn't need to perpetrate tally fraud; he needed to defend his real tallies against false accusations of tampering. And the system worked remarkably well for that purpose.

Prior to 1998, Venezuelans cast votes via paper ballots that were counted by hand. In the elections of 1998, 1999, and 2000, the electoral council gradually introduced paper ballots that were counted by machines. Then, beginning with the recall referendum of August, 2004, the Chávez administration introduced a new form of electronic voting. While not identical to the system used in 2024, which I describe in the previous section, Venezuela's 2004 electronic voting system included many of the same key features: a screen-based ballot featuring candidate photos and party colors; an individual paper ballot receipt for each voter, reporting her vote choice; a de jure mandatory hand count of these ballot receipts in a sample of ballot boxes; machine-printed tally sheets, with copies retained by each witness; and the requirement that the government publish voting-machine-level tallies (to be compared with the paper tally sheets) within thirty days.

In the two decades since the installation of Venezuela's electronic voting system, it has validated disputed government-published tallies on many occasions and, on other occasions, provided clear proof of government tally fraud.

Example of a case in which the electronic voting system validated disputed official tallies. In Venezuela's 2013 presidential election, the voting-machine-level

tallies published by the Venezuelan electoral council indicate that Maduro won that election with a narrow margin of 1.5 percentage points (50.61% to 49.12%), or approximately 224,000 of more than 15 million votes.

The Capriles campaign challenged the results, claiming fraud. Even though the hot audit (i.e., the hand count of paper ballots in more than half of ballot boxes) all but ruled out systematic rigging of voting machines, meaning that the machines themselves functioned as expected (Carter Center, 2013), the Capriles campaign initially requested a "recount" (as the Carter Center pointed out, the meaning of recount is unclear in a context in which electronic votes are tallied by software). The electoral council responded to Capriles's request by counting all paper ballots by hand (i.e., expanding the audit from the original 53% of ballot boxes to 100%, BBC 2013). The hand-counted paper ballot receipts matched the machine tallies, save for a minuscule number of missing ballot receipts attributed to voters pocketing their receipts (whether accidentally or in order to prove that they voted for the government in the carrusel, Clímax 2017).

Capriles also denounced voter impersonation: individual voters arriving at multiple voting booths with fake ID cards, obtaining manual overrides in order to cast votes when their fingerprints were found not to match the fingerprint database. The press documented several such cases (La Patilla, 2013), and the electoral council's subsequent duplicate fingerprint audit—in which technicians combined automated and human comparisons of fingerprints across all voting booths, according to the electoral council—uncovered 245 definite cases of duplicate fingerprints and an additional 10,726 cases of possible duplicates (Carter Center, 2013, 73–74). Even if all possible duplicates were indeed duplicates, therefore, and even if all voter impersonation favored Maduro, this form of voter fraud would not have been close to sufficient to change the outcome of the election.

The fingerprint audit did not satisfy Capriles or the Mesa de Unidad Democrática (MUD), neither of whom recognized Maduro's victory. Moreover, it remains unclear whether María Corina Machado believes that Venezuelan voting machines counted votes correctly in 2013 or in other prior elections; in an interview on August 17, she said, "For the first time in 25 years and after dozens of fraudulent elections, we have the proof" (El Mundo, 2024). But governments all across Latin America

and Europe did recognize the outcome of the 2013 election (McCarthy and McCoy, 2013), and the United States quickly resumed dialogue with Maduro (El País, 2013). Moreover, the MUD soon afterward mobilized to compete in the 2015 legislative elections, suggesting considerable internal confidence in vote-counting. In other words, Venezuela's electronic voting system convinced substantial internal and international audiences that more voters cast ballots for Maduro than for Capriles in a very close and hotly contested election.

Of course, the electronic voting system in no way guards against intimidation of opposition voters, coerced mobilization of pro-government voters, or use of government resources for the campaign, all of which occurred in the run-up to Venezuela's 2013 presidential election (Carter Center, 2013). I focus on the validity of vote-counting itself not because these other forms of manipulation are unimportant but because they are less relevant to the objective of this note: assessing competing claims about the outcome of Venezuela's 2024 presidential election.

Examples of cases in which the electronic voting system revealed government tally fraud. Just as the paper trail from Venezuela's electronic voting system established the truth of no tally fraud in dozens of Venezuelan elections since 2004, it also quickly revealed tally-fraud attempts in other instances. In the gubernatorial election in the state of Bolívar in 2017, for example, the voting-machine-level tallies posted on the website of the electoral council did not match the paper tally sheets printed on site and retained by opposition witnesses (Toro, 2017). And in the gubernatorial election in Barinas in 2021, in which the MUD candidate Freddy Superlano ran against Argenis Chávez (Hugo's brother), the Superlano campaign collected tally sheets establishing that Superlano won. Venezuela's national electoral council credited the campaign tally sheets and looked to be moving toward declaring him the winner, only to have the Supreme Court declare Superlano ineligible for public office (Rodríguez, 2024a). In that instance, as in Bolívar in 2017, Venezuela's electronic voting system revealed government attempts to steal gubernatorial elections.

In the 2023 referendum on Venezuela's claim to the territory disputed with Guyana, and in the election of members to a Constituent National Assembly in 2017—both of which the opposition boycotted—the government broke with tradition (and, in

the case of the assembly, with regulation)⁶ by not publishing voting-machine-level tallies. This absence of disaggregated data strongly suggests that the announced top-line results did not reflect votes cast (as in the 2024 presidential election considered here).

The special case of 2004. Venezuela's electronic voting system was introduced in the 2004 recall referendum, in which Venezuelans cast up-or-down votes on whether to recall Hugo Chávez. When opposition leaders began mobilizing to collect signatures for a petition requesting the referendum, Chávez's popularity ratings were low; had the referendum been held promptly, Chávez might well have been recalled. But the government successfully delayed the referendum; by the time it was held, in August of 2004, Chávez's approval ratings had skyrocketed, fueled by the suddenly rising price of oil. Public opinion polls predicted that voters would choose to retain Chávez by a large margin, and the official tallies indicated that 59% voted to retain him. In part because two exit polls indicated that voters had chosen to recall Chávez by a large margin, politicians and activists claimed that the voting machines had been rigged (McCoy and Diez, 2011).

Many of the key features of Venezuela's electronic system in 2004 were as they are today: each voter received a paper ballot receipt and deposited it in a ballot box; each voting machine printed a paper tally sheet at the end of the day, with copies for opposition witnesses; and the electoral council promptly published voting-machine-level tallies, which turned out to match the machine-printed tally sheets.

But there was one important difference. In 2004, the electoral council did not require poll workers to randomly select half of ballot boxes in each polling place to be opened for hand-counting of paper ballot receipts. Instead, the electoral council selected a sample of just 1% of ballot boxes nationwide, and informed poll workers at the selected voting booths of their responsibility for hand-counting. This sample turned out not to be randomly selected (Hausmann and Rigobon, 2011), meaning that the hand-count audit alone cannot rule out the possibility that the electoral

⁶Article 382 of the *Reglamento General de la Ley Orgánica de Procesos Electorales* requires the electoral council to publish voting-booth-level tallies "for executive positions and for deliberative bodies," but does not mention referenda. Still, the electoral council did publish voting-booth-level tallies for referenda in 2004, 2007, and 2009.

council rigged the other 99% of not-audited machines. At the request of international observers, the electoral council did conduct a second hand-count audit of another sample of ballot boxes three days after the election, but the chain of custody of materials was not sufficient to rule out tampering with the paper receipts themselves.

It was because of sustained controversy over the 2004 election that the electoral council later strengthened the hand-count audit, requiring poll workers to open approximately 50% of boxes nationwide (rather than 1%), and allowing poll workers to locally select which boxes to open (according to idiosyncratic on-site randomization procedures like drawing from a hat) rather than drawing the sample centrally. This strengthening of the hand-count audit protected future elections from the kind of persistent doubt that clings to 2004.

Summary. This is all to say that the double paper trail in Venezuela's electronic voting system has historically provided robust evidence about what happens on election day. Time and time again, the paper trail has either confirmed the validity of voting-machine-level tallies published by the electoral council (as in 2013) or revealed that those tallies are false (as in gubernatorial elections in Barinas and Bolívar). It is this paper trail that allows researchers to treat Venezuelan electoral returns from this period as meaningful measures of votes cast (e.g. Albertus, 2015; Kronick et al., 2023). This history provides every reason to expect that the paper trail can now either confirm or refute the validity of the voting-machine-level tallies published by the campaign of Edmundo González Urrutia.

3 The campaign data

Anticipating that the Venezuelan electoral council might not publish voting-machine-level tallies, or might publish fabricated voting-machine-level tallies, the González campaign organized hundreds of thousands of volunteers to help collect and scan tally sheets on election night (Rogero, 2024). Campaign witnesses were meant to retain copies of the printed tally sheets from each voting machine (as is their right under Venezuelan law), take a photo of the QR code at the bottom and transmit both the photo and the QR-code-produced data through the campaign's dedicated app, and then also bring the printed tally sheet to a local campaign command to

be scanned. Within two days, the campaign had posted the tally-sheet scans on a website and distributed a spreadsheet organizing the results.

As of this writing, the campaign data include vote counts from 25,073 voting machines, 82.8% of the total of 30,280 voting booths; this sample includes 84.7% of registered voters. In the sample, Edmundo González obtained 67% of valid votes. Even if Maduro had obtained 100% of votes from all registered voters (i.e., 100% turnout) in the remaining 17% of voting machines, González would still have won with a margin of five points (51.5% to 46.5%).

4 The paper trail in the 2024 presidential election

Ideally, the paper trail would definitively answer two questions: (1) Was the campaign data extracted from the voting-machine-level tally sheets that were printed on site (i.e., at each polling place) on election day? And, (2) If so, do these tally sheets reflect voters' actual choices? If the answer to both of these questions were *yes*, then we would conclude that the campaign data accurately capture votes cast.

The Associated Press (2024) and The Washington Post (2024) independently compared the tally-sheet *images* published by the campaign to the data in the spreadsheet published by the campaign, finding that they match. This analysis establishes that the campaign data were indeed extracted from a set of paper tally sheets (rather than, say, fabricated wholesale), though it does not establish the *tally sheets printed* on site on election day as the source. If the campaign had prior or subsequent access to the voting machines, staffers could have printed a separate set of tally sheets (i.e., a set different from those printed on site on election day).

But in practice, had this occurred, a government witness would almost certainly

⁷In the campaign data, there are 10,887,262 valid votes, of which Maduro won 3,316,142 and González won 7,303,480. There are 3,269,116 voters registered at the remaining voting booths (i.e., those not in the campaign data). If, improbably, Maduro won all of these votes, he would have 6,585,258 of a total of 14,156,378 valid votes, or 46.5%, to González's 51.5%. For that reason, estimating Maduro's vote share in the remaining (i.e. missing) voting booths is not necessary in order to establish the outcome (conditional on valid campaign data). But the overall vote shares are politically relevant, in the sense that González+5 implies different political possibilities than González+35; for that reason, I propose an approach to estimating the overall vote shares in Appendix B; this approach suggests that González's overall vote share was approximately 66%.

have noted and denounced the discrepancy. Just as González campaign witnesses were entitled to retain a copy of the printed tally sheets, Maduro's witnesses were also so entitled. And anyone in possession of a tally sheet with totals different from those in the campaign data would face strong incentives to say so. At this writing, to the best of my knowledge, there are no credible reports of rival tally sheets (one recent report of a rival tally sheet was quickly debunked). Moreover, journalists have published a collection of videos that appear to show poll workers reading the tally sheets to the public on election night; the numbers in these videos match those in the campaign-published tally sheets (Cazadores, 2024).

That the campaign data were extracted from paper tally sheets printed by the voting machines on site on election day does not, in and of itself, establish that those tally sheets reflect votes cast. In theory, the voting machines could be pre-programmed to (a) print the correct ballot receipt for each voter even while (b) printing tally sheets that do not reflect sums of votes cast. In fact, this is the very form of fraud of which critics accused the government in the 2004 recall election (see previous section). Voting machines could (to take a naïve example) flip every fifth Maduro vote in favor of González for the purposes of the tally sheet, without revealing this malfeasance directly to individual voters in the form of incorrect ballot receipts. But as noted above, Venezuela's electronic voting system is designed to prevent this type of tally fraud by requiring that poll workers open and count the paper ballots in at least half of ballot boxes, in the presence of all candidates' witnesses, filling out a pen-and-paper form with the totals (Figure 1d).

While witnesses are *de jure* allowed to retain copies of the machine-printed tally sheets, and typically do so in practice, they often do not retain a carbon copy of the pen-and-paper hand-count verification form—and the González campaign did not systematically collect photos of these forms through the campaign smartphone app. When campaign officials asked witnesses (at my request) if they had happened to take photos of the hand-count forms, several dozen did share such photos; all of the legible ones do match the numbers of the corresponding tally sheets.

But more revealingly, government officials do retain the hand-count verification forms, returning them to the electoral council. Presumably, if the hand-counted paper ballots reflected more votes for Maduro than appear on the printed tally sheets

from which the campaign data were extracted, the government would have every incentive to publicize this fact. At this writing, to the best of my knowledge, there are no reports of such discrepancies.

This point bears repeating: Venezuela's electronic voting machines produce a paper ballot receipt for each voter. Those receipts are deposited in ballot boxes (one for each machine). At the close of voting, poll workers are required to randomly select at least half of ballot boxes in the precinct (see Footnote 4), open those boxes, count the tickets in the presence of all party witnesses, and confirm that the hand-counted total matches the machine-printed tally sheet. Eugenio Martínez, a prominent journalist who has long covered Venezuelan elections, told me that the proportion of boxes actually opened is likely closer to 30% than the prescribed 55%. But the point stands. Venezuelan voting machines are "just the world most expensive pencil: they produce printouts, and the printouts are counted by hand. No hacker in Timbuktu or anywhere else can do anything about that" (Toro, 2015).

In order to believe that the campaign data do not reflect the votes cast on July 28, 2024, one would therefore have to think that: (1) The González campaign was able to work with or hack voting-machine vendor Ex-Cle in order to rig the software to flip votes in favor of their candidate for the purposes of the tally sheets, without issuing inaccurate ballot receipts to individual voters, and that (2) not a single pro-Maduro witness nor a single official of the electoral council has chosen to publicize a single one of the (likely) thousands of hand-count verification sheets that would provide clear evidence of such malfeasance, had it occurred. A campaign-Ex-Cle collaboration strikes me personally as implausible, given that the electoral council chose and vetted Ex-Cle as a vendor, and given reports of the nature of the relationship between Ex-Cle and the Venezuelan government (ArmandoInfo, 2024). The silence of thousands of pro-government witnesses and many government officials in possession of evidence of opposition fraud strikes me as yet more implausible. Of course, each person may judge for herself the likelihood of this alternative hypothesis; my objective is merely to establish what would have had to occur, given the paper trail, for the campaign data to misrepresent votes cast.

Imagine that the González campaign *did* successfully collaborate with or hack Ex-Cle, and that thousands of people in possession of paper evidence of this fraudincluding government officials—have chosen to remain silent. In that case, the machines might have (for example) reassigned a certain proportion of Maduro votes to González, where the proportion is a sensible function of location and/or the history of vote shares in that voting machine. This type of fraud would not produce anomalous patterns in the final digits of vote counts, or suspicious quantities of duplicate values, or odd deviations from political geography in previous elections. For that reason, while it will be valuable on substantive grounds to merge the campaign data with that of past elections and to investigate how González built such a large coalition, such exercises are not especially informative about whether the campaign data derive from universally rigged machines.⁸

Yet one more possibility, however remote, is that the campaign somehow divined which ballot boxes would be opened (perhaps by suggesting a specific box-selection mechanism to opposition-aligned voters) and rigged only those voting machines associated with ballot boxes that would not be audited. But the electoral forensics analysis in Mebane (2024) would pick up selective malfeasance of this sort. Instead, consistent with the conclusions of my analysis of the paper trail, Mebane (2024) finds that the campaign data are not fraudulent.

5 In the event of dueling tally sheets

At this writing, it appears possible that the Venezuelan government will publish voting-machine-level data that adds up to the government's announced vote totals, perhaps along with a set of tally sheets that match the disaggregated data. If the government were to fabricate this data, it might do so in a clumsy and obvious way (as an extreme hypothetical example, it could publish data indicating that Maduro won the same share of the vote in every polling place), but it could also without much difficulty fabricate data using methods undetectable by quantitative forensic analysis.

This potential dueling-tally-sheets scenario is one in which the hand-count verification forms would once again prove useful. Unlike the tally sheets or even the individual paper ballot receipts, these hand-count verification forms are not printed

⁸I thank Uri Simonsohn for convincing me of this point.

by a machine; poll workers fill them out with pen and they are signed by all witnesses, making them somewhat difficult (or at least time-consuming) to forge. That is why observers interested in adjudicating between two rival sets of tally sheets should request that the electoral council publish not only voting-machine-level data—which Brazil and others have already requested—but also the hand-count verification forms.

6 Conclusion

For many observers, the notion that the campaign data are real—and that therefore Edmundo González Urrutia won the 2024 presidential election in a landslide—might seem so apparent as to obviate so much exposition. Yet many others, especially but not exclusively observers outside Venezuela, begin from the premise that we should not underestimate the organizational prowess, conspiratorial acumen, or financial resources of politicians who oppose Chavismo. The evidence presented here indicates that even extraordinary levels of organizational prowess, conspiratorial acumen, and resources could not perpetrate tally fraud that would produce the campaign data without leaving traces in the paper trail—traces that, at least as of this writing, have not appeared. For that reason, even a person inclined to doubt the Venezuelan opposition may also, without any conflict, believe that the campaign data reflect the ballots cast by Venezuelan voters on July 28, 2024.

References

- Albertus, M. (2015). The role of subnational politicians in distributive politics: Political bias in venezuela's land reform under chávez. *Comparative Political Studies*, 48(13):1667–1710.
- ArmandoInfo (2024). Cómo privatizar unas elecciones. Armando.info.
- Associated Press (2024). Ap review of venezuela opposition-provided vote tallies casts doubt on government's election results. By Regina Garcia Cano, Joshua Goodman, and Angeliki Kastanis.
- BBC (2013). Venezuela: Hay que recontar los votos por la paz del país. BBC News Mundo.
- Brewer-Carías, A. R. (2010). Dismantling democracy in Venezuela: The Chávez authoritarian experiment. Cambridge University Press.
- Carter Center (2005). Observing the venezuela presidential recall referendum: Comprehensive report. The Carter Center.
- Carter Center (2013). Study mission of the carter center 2013 presidential elections in venezuela. The Carter Center.
- Cazadores (2024). Euforia interrumpida: Así coinciden los videos y fotos tomados el 28j con los resultados publicados en línea por la oposición. Cazadores de fake news.
- Clímax (2017). El "carrusel" del psuv: estocada final al voto secreto. By Andrea Tosta.
- CNE (2024). Manual de funcionamiento genérico de la mesa electoral. Consejo Nacional Electoral de Venezuela.
- Corrales, J. (2020). Democratic backsliding through electoral irregularities. European Review of Latin American and Caribbean Studies/Revista Europea de Estudios Latinoamericanos y del Caribe, (109):41–65.
- Corrales, J. and Penfold, M. (2011). Dragon in the Tropics: Venezuela and the Legacy of Hugo Chavez. Brookings Institution Press.
- Counterpunch (2024). Venezuela: An attempted coup by any other name. By Maria Paez Victor.

- Delfino, G. and Salas, G. (2011). Analysis of the 2004 venezuela referendum: The official results versus the petition signatures. *Statistical Science*, pages 479–501.
- Democracy Now! (2024). Venezuela: Maduro claims victory, accuses opposition of coup attempt following disputed election. democracynow.org.
- El Mundo (2024). María corina machado: Aquí no habrá reparto de poder ni nuevas elecciones. By Cayetana Álvarez de Toledo.
- El País (2013). Us and venezuela agree to "move forward" to fix soured bilateral ties. By Pablo Ximénez de Sandoval.
- Ellner, S. (2024). What the mainstream media isn't saying about venezuela's maría corina machado. NACLA.
- Fujiwara, T. (2015). Voting technology, political responsiveness, and infant health: Evidence from brazil. *Econometrica*, 83(2):423–464.
- Gamboa, L. (2022). Opposition at the Margins. Cambridge University Press.
- Goel, S., Meredith, M., Morse, M., Rothschild, D., and Shirani-Mehr, H. (2020). One person, one vote: Estimating the prevalence of double voting in us presidential elections. *American Political Science Review*, 114(2):456–469.
- Gulotty, R. and Kronick, D. (2022). The arbitrage lobby: Theory and evidence on dual exchange rates. *International Organization*, 76(1):105–125.
- Hausmann, R. and Rigobon, R. (2011). In Search of the Black Swan: Analysis of the Statistical Evidence of Electoral Fraud in Venezuela. *Statistical Science*, 26(4):543–563.
- Hawkins, K. A. (2010). Venezuela's Chavismo and populism in comparative perspective. Cambridge University Press.
- Kronick, D., Plunkett, B., and Rodriguez, P. L. (2023). Backsliding by surprise: the rise of chavismo. *Political Science Research and Methods*, 11(4):838–854.
- La Patilla (2013). Diputado oficialista acompaña a hombre con 40 cédulas laminadas. Lapatilla.com.
- McCarthy, M. and McCoy, J. (2013). The limits of legacy: The post-chávez challenge and electoral legitimacy. Americas Quarterly.

- McCoy, J. and Diez, F. (2011). *International Mediation in Venezuela*. United States Institute of Peace.
- Mebane, W. (2024). eforensics analysis of the venezuela 2024 presidential elections. Working Paper.
- Rodríguez, F. (2008). An empty revolution-the unfulfilled promises of hugo chavez. Foreign Aff., 87:49.
- Rodríguez, F. (2024a). How clientelism works: Evidence from the barinas special election. World Development, 184.
- Rodríguez, F. (2024b). Scorched Earth. Notre Dame University Press.
- Rodríguez, F. (2024c). Sobrepredicción del voto opositor. X.
- Rogero, T. (2024). How venezuela's opposition proved its election win: 'a brilliant political move'. The Guardian.
- The New York Times (2024). U.s. recognizes maduro's rival as winner of venezuelan election. By Julie Turkewitz.
- The Wall Street Journal (2024). U.s. says opponent of venezuela's maduro won presidential election. By Alan Cullison.
- The Washington Post (2024). Maduro lost election, tallies collected by venezuela's opposition show. By Samantha Schmidt, Steven Rich, Ana Vanessa Herrero and María Luisa Paúl.
- Times, T. N. Y. (2024). Venezuela's strongman was confident of victory. then came the shock. Anatoly Kurmaneav.
- Toro, F. (2013). Slow and steady wins the race. Caracas Chronicles.
- Toro, F. (2015). Venezuela's election: The basics. Caracas Chronicles.
- Toro, F. (2017). PSUV Steals Bolívar State Governor's Race. Caracas Chronicles.

Appendix

A	Additional discussion of the possibility of fraud in 2004	20
В	An estimate of overall vote shares	22

A Additional discussion of the possibility of fraud in 2004

In the main text, I note that controversy over the 2004 recall referendum illustrates the importance of a convincing hand-count audit of paper ballot receipts. This is the first and principal reason that the experience of 2004 is relevant to discussion of the 2024 presidential election: the insufficiency of the hand-count audit in 2004 was the impetus for the strong hand-count audit that is in place today.

In this appendix, I reconsider claims that there is positive evidence of widespread, voting-machine-rigging fraud in the 2004 recall referendum. Determining whether or not there is positive evidence of fraud in 2004 is not essential to the central task of this research note, which is to evaluate the validity of the data published by the González campaign. Yet it is relevant to the political context more broadly.

Any allegation of fraud in 2004 must accommodate three facts. First: three days after the election, ballot receipts from a sample of ballot boxes were hand-counted in the presence of international observers (Carter Center, 2005, 88). Whether the sample was random is disputed, but it is clear that the sample is politically representative, meaning that prior political outcomes in the sample match those outcomes in the universe of electronic-voting booths. Second: the ballot receipts from that sample of boxes add up to the machine-tallied totals announced by the electoral council. And third, the mean recall-referendum result is the same in the audited sample as in the universe of electronic-voting booths.

A coherent theory of fraud that can accommodate these three facts is necessarily somewhat complicated, but I believe that such a theory exists Delfino and Salas (2011). The electoral council could have: (1) received the actual results on election night from the electronic voting machines; (2) quickly analyzed the relationship between the share of pro-recall votes (si votes) and the share of voters who had signed the petition requesting the recall; (3) subset the data to those voting machines with especially low votes-to-signatures ratios, i.e., those in which petition signatures produced relatively few pro-recall votes (call this the low-yield subset, i.e., places that underperformed from the perspective of the opposition); (4) selected a random sample from that subset and exempted it from tampering, for the purposes of the subsequent hand-count audit; and then (5) fabricated results for the remaining voting machines, calculating the false pro-recall vote share as a linear function of the signatories share (plus noise), using the slope in the low-yield subset. This

⁹The planned hot audit, on election day, did not take place.

¹⁰Specifically, vote shares in 1998 and the fraction of voters who had signed the petition requesting the recall referendum.

method would require the electoral council to be able to transmit results to the voting machines prior to the printing of tally sheets, rather than merely receiving results transmitted from the voting machines; it appears that such two-way communication did in fact take place (Hausmann and Rigobon, 2011, 545).

To illustrate this theory, imagine that each petition signature actually corresponded to 1.5 pro-recall votes, on average, but produced only one pro-recall vote in the low-yield subset. The electoral council could have selected voting machines from the low-yield subset for the hand-count audit and then forced the remaining voting machines to follow the one-signature-one-vote relationship.

As evidence in favor of this hypothesis, Delfino and Salas (2011) note that the correlation between (a) the recall-Chávez vote share and (b) the proportion of voters who had previously signed a petition requesting the recall referendum was much higher across electronic voting machines ($\rho = 0.98$) than across the small subset of voting booths that retained paper ballots ($\rho = 0.6$). This evidence appears compelling: why should the correlation between votes and signatures be so much tighter across voting machines than across paper-ballot voting booths, if not because the government programmed the voting machines to calculate votes as a linear function of petition signatures (plus noise)? As Delfino and Salas note, the number of pro-recall votes "seems to behave in an excessively linear fashion relative to the number of signatures in support of the recall referendum" (493).

But in fact, paper-ballot voting booths differed from booths with electronic voting machines in two ways that should affect the strength of the signatures-votes correlation. First, all embassies and consulates used paper ballots, and voters abroad had not enjoyed the same opportunity to sign the recall petition as voters located within Venezuela. Second, paper-ballot voting booths were typically smaller than those using electronic voting machines: the average both with electronic voting had 2,700 registered voters, compared to just 441 for paper-ballot voting booths. Simply excluding embassies and consulates and excluding voting booths with more than 600 petition signatures (i.e., restricting to smaller voting booths), makes the signatures-votes correlation equal: $\rho = 0.933$ in booths with electronic voting, $\rho = 0.932$ in booths with paper ballots.

B An estimate of overall vote shares

The campaign data includes enough voting booths (mesas) and indicates a sufficiently large González lead (37 percentage points) that it excludes the possibility of a Maduro victory, even if he were to win 100% of votes from 100% of registered voters in all remaining voting booths (again, if the campaign data are accurate). For that reason, estimating González's vote share in the remaining voting booths (i.e., those left out of the campaign data) is not necessary in order to establish the outcome of the election.

But the overall vote shares are politically relevant for other reasons. In the implausible scenario in which Maduro did in fact obtain 100% of votes from 100% of registered voters in all remaining voting booths, González would have won, 51.5% to 46.5%; at the other extreme, if González had obtained 100% of votes from 100% of registered voters in all remaining voting booths, he would have won with a much larger margin: 74.7% to 23.4%. Where the actual outcome lies between these two implausible extremes affects how Maduro, González, and their respective co-partisans might form expectations about the outcome of future elections, were there to be any.

Empirical approach. The campaign data include electoral returns for 25,073 of 30,027 voting booths.¹¹ My objective is to estimate both turnout and González's vote share in the remaining 4,954 voting booths, using each voting booth's electoral history. I train two random forest models on the 25,073 voting booths in the campaign data, one in which the outcome is 2024 turnout, and another in which the outcome is González vote share; the predictors are electoral outcomes from 2012, 2013, 2015, and 2018, as well as state fixed effects.¹² Because these predictors are strongly related to electoral outcomes in 2024, the models achieve low (out-of-bag) mean squared error within the campaign data, suggesting that they likely also make good predictions for the voting booths outside the campaign data.¹³

¹¹The count of 30,027 excludes 195 in embassies and consulates as well as 59 very small voting booths in hard-to-reach locations; I omit these 254 from my analysis.

¹²Specifically, I include Capriles's vote share in 2012; turnout in 2012; Capriles's vote share in 2013; turnout in 2013; MUD's vote share and PSUV's vote share in the 2015 voto lista congressional election (unlike the 2012 and 2013 presidentials, 2015 was not a two-way race, making it meaningful to include both); turnout in 2015; and, for 2018, the share of registered voters who did not cast votes for Maduro (i.e., abstainers and opposition voters as a share of the electorate). I use the same predictors for both outcomes.

¹³The random-forest predictions for the 4,954 missing voting booths are strongly correlated ($\rho = 0.93$) with predictions from a simple regression in which electoral history enters linearly.

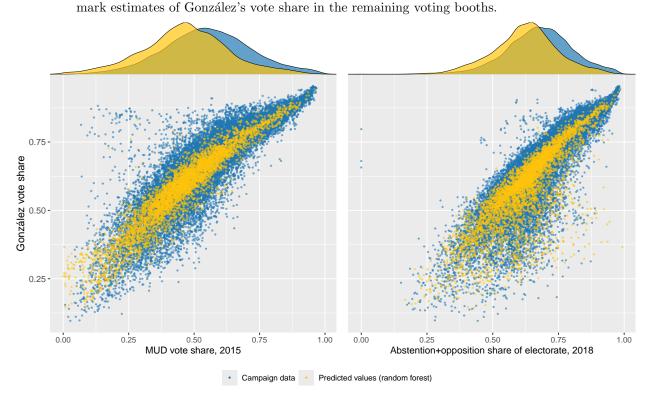
Data. Voting-booth-level data for the 2024 election were published by the González campaign at https://resultadosconvzla.com/; for previous elections, voting-booth-level data were published online by the CNE (the electoral council).

Merging 2024 voting booths to voting booths in prior elections is not entirely straightforward, for two reasons. First, there are voting booths—and indeed entire precincts (centros de votación)—that existed in 2024 but did not exist in a given previous election. Of the 30,027 voting booths in 2024, for example, only 26,290 (88%) existed in 2018. Second, and relatedly, even a voting booth that has the same ID number as its predecessor is not exactly the same. Two voting booths with the same ID number are largely the same: precinct ID numbers are stable across elections, meaning that the same numeric ID indicates a precinct in the same location; within precincts, voter assignment to booths is also largely stable (voters are assigned according to the final two digits of their cédula numbers). But of course, some voters move, and when a precinct grows or shrinks—or a new one opens or closes nearby—some voters are reassigned.

To deal with both of these issues, I use the electoral registry to calculate the proportion of voters in each 2024 voting booth who were assigned to that same booth in the previous election of interest (this proportion is zero for new voting booths), as well as the proportions of voters who migrated from each voting booth that existed in the previous election. For a given voting booth in 2024, for example, I might observe that 85% of registered voters were registered in that same booth in 2015; 10% migrated from one nearby voting booth, 4% migrated from a second nearby voting booth, and 1% were newly registered. I would then estimate this voting booth's electoral history—its 2015 MUD vote share, for example—as a weighted sum of vote shares in the three parent voting booths (weighted by the proportion of voters who originate in each parent mesa). One advantage of this approach over a naïve merge on voting-booth ID number is that it does not require me to drop new voting booths from the analysis; another is that it takes into account, if imperfectly, how existing voting booths change over time. As it happens, the top-line results that I obtain via this method are quite similar to the results from a naïve merge on ID numbers.

Results. The campaign data indicate that González obtained 67% of the vote in the 25,073 voting booths for which the campaign has tally sheets; the random forest model predicts that he obtained 60% of the votes in the remaining 4,954 voting booths. Given that turnout was 60% in the campaign-data voting booths, and given the model estimate of 62% turnout in the remaining voting booths, and taking into account the number of registered voters at all voting booths (which is observed, not

Figure B.1: González vote shares in the campaign data and elsewhere Blue points mark González's vote shares in each voting booth in the campaign data, plotted against MUD's vote share in 2015 (left panel) and the proportion of registered voters who did not cast ballots for Maduro in 2018 (abstainers + opposition voters). Yellow points



estimated), these estimates suggest that González's overall vote share was 66%.

Figure B.1 plots the campaign-data tallies (blue) and predicted values (yellow) of González's vote share against the 2015 MUD vote share (left panel) and against the share of registered voters who did not vote for Maduro (that is, abstainers and opposition voters together as a share of the electorate) in the 2018 presidential election (which the MUD boycotted). The marginal density plots show the distribution of prior outcomes (2015 and 2018) in voting booths that later entered the campaign data (blue) and voting booths that did not enter the campaign data (yellow). These distributions show that, while campaign-data voting booths do lean more opposition (as many observers have supposed), the distributions are not entirely dissimilar: in 2015, the MUD obtained 55% of votes in the voting booths that ended up in the campaign data, compared to 48% of votes in the remaining voting booths. It is not

surprising, then, that the model also estimates a seven-percentage-point difference between these two groups in 2024.