Updated Comment on the 2019 Bolivia Presidential Election and OAS Statistical Analysis

Irfan Nooruddin

26 August 2020

This note updates a previous note written on 19 August 2020. (Some of my response to earlier critiques is in that note, so, for completeness, and to preserve the record, that note is appended to the end of this one as Appendix A.)

Below I describe in detail the modeling choices made in my analysis of the 2019 Bolivian Presidential Election results that were included in the Final OAS Report, which is available at https://www.oas.org/fpdb/press/Audit-Report-EN-vFINAL.pdf (specifically pages 86-93).

The code and data archived in the Harvard Dataverse replicates every figure and number reported.

In particular, I address three specific criticisms that have been raised against my analysis:

1) *Omission of 4.1% polling stations in the analysis of the TREP data set*: I explain below the reasoning behind this choice, explain that these polling stations were included in the tabular analysis presented, and show that including them does not alter the finding reported.

2) *Choice of 'statistical estimator' used to draw the trendlines to identify a break*: Idrobo et al. (2020) argue that I should have used a local linear regression instead of a local means smoother. This is a valid point and I agree with them. Using the local linear regression estimator yields the same results. For ease of exposition, all models below therefore use the local linear regression technique. I thank Idrobo et al. for this suggestion.

3) *Mistake in the Computo timestamp*: Rosnick (2020) identifies a mistake made in how the Computo timestamps were calculated. I am grateful for this correction. I have updated the archived data set with the corrected timestamp in my Dataverse, and replicate all affected results using the corrected data series. None of the results change and I stand by all previously reported conclusions.

In the remainder of this note, I provide a step-by-step explanation of all choices made. I conclude with my assessment of the analysis conducted.

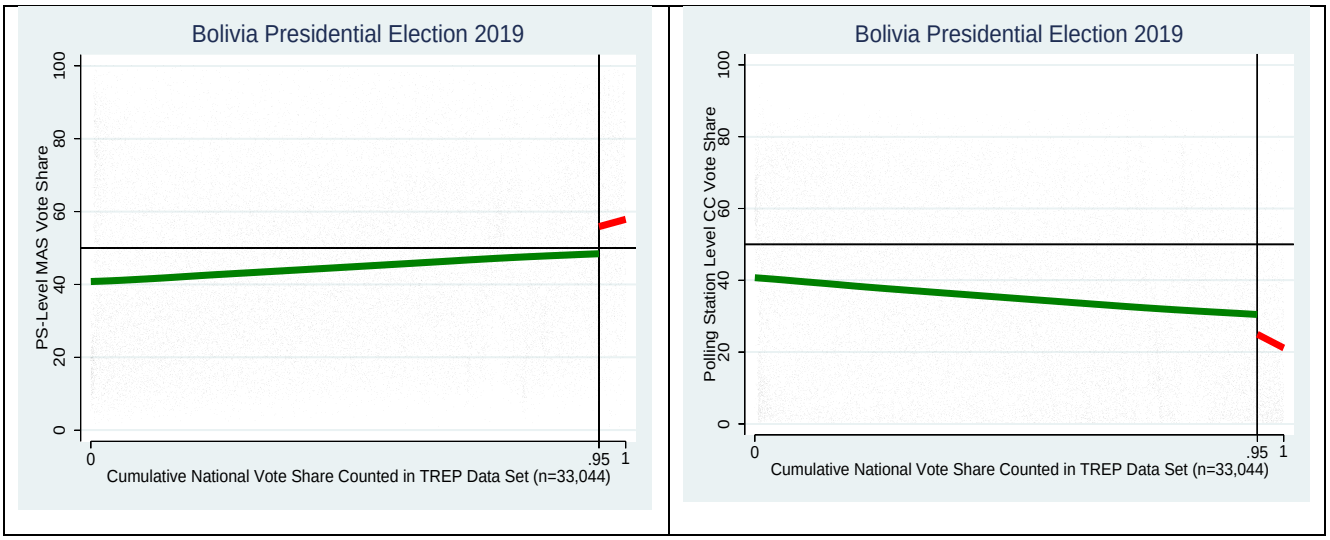
ANALYSIS OF THE TREP DATASET

The analyses presented on pages 88-91 utilize the TREP data provided by the Bolivian TSE. This data set provided the vote tallies for 33,044 polling stations; 1,511 polling stations were not included in this data set by the TSE.

The key facts were that with 84% of the votes tallied in the TREP data, the MAS candidate, incumbent President Evo Morales, was below the 10% vote margin threshold required to avoid a second round run-off election. By the time the TREP count was complete, he had cleared that threshold. I focused my attention therefore on that last period of the vote count.

The TREP data set includes a timestamp recorded by the TSE for when the *acta* was received and verified (*verificador_date*). This allows us to put the polling stations in the order that their *actas* were verified by the TSE. This is the same technique used by CEPR in their October 2019 report. Because 1,511 polling stations were not included in the TREP data set, these polling stations do not have a timestamp in the TREP data set.

I began my analysis by looking to see if there were any breaks in the vote trends at the polling station level. This led me to find a break in the trends for the two main contenders around the 95% cumulative vote count mark of the TREP data set. Using the local linear regression estimator suggested by Idrobo et al. (2020), the figures for the MAS (on the left) and Civic Community (on the right) are reproduced below.



While the MAS polling station level vote share had been trending upwards throughout the count, and the CC's trending downwards throughout the count, there appears to be a "break" in the trendlines, which I suggested warranted explanation. Note that I never suggested that this was definitive proof of irregularities. Nor is it. It is a diagnostic tool, and using it the OAS field teams investigated further, and identified a set of facts that they describe in detail in their report. It is those facts that led them to issue the report they did; if the graphs above played any role, it was by helping the OAS focus its attention on a set of polling stations. But had the further investigation into those polling stations yielded nothing of concern, the OAS would have not issued the report it did.

(As an aside, I find puzzling the attention to these figures. There was never a claim that these figures were "proof" of fraud nor does the word 'fraud' appear in pages 86-92 of the OAS report which was based on my analysis. These figures were a diagnostic tool that focused the attention of the OAS field teams, whose on-ground analysis drove their final conclusions. Obsessing about the graphs is akin to worrying whether a thermometer was perfectly calibrated even after the suspicion of a fever leads one to go to a doctor who, after further testing, diagnoses covid-19.)

What about the 1,511 polling stations not included in the TREP data set? Idrobo et al. (2020) suggest that these were incorrectly excluded above. This is not the case. The figures above

utilize the TREP time-stamps, which those polling stations do not have. Therefore, there is no straightforward way to include them though they were included in all the tables presented in page 90. Specifically, lacking any basis by which to know the time at which these 1,511 polling stations reported their votes, I treated them as "late reporters", which I stated transparently and explicitly on page 86 of the OAS final report:

> "1,511 polling stations were not included in the TREP but do appear in the final Computo results, which are the official vote tallies of the Bolivian system. All the analysis conducted below include these additional polling stations. Since they were not included in the TREP, they are treated as being late reporters. We stress that all the results below are based on the Computo vote tallies."

Consider the second table on page 90, where I explicitly separate the "Computo-only" polling stations[1]:

	PS-level MAS Vote Share			PS-level CC Vote Share		
	0-95% of TREP	95-100% of TREP	Computo only	0-95% of TREP	95-100% of TREP	Computo only
National	44.6	56.9	49.6	35.4	23.0	28.6
Beni	31.5	45.9	43.7	34.1	22.2	24.6
Chuquisaca	40.7	58.6	66.8	41.0	19.9	14.6
Cochabamba	54.0	65.4	62.3	33.1	21.5	24.7
La Paz	50.5	56.8	60.3	29.0	18.5	17.9
Potosi	47.3	55.3	51.3	29.8	22.7	25.3
Santa Cruz	33.1	40.6	37.1	45.1	33.3	39.8
Tarija	38.8	41.7	32.3	38.9	35.3	46.9

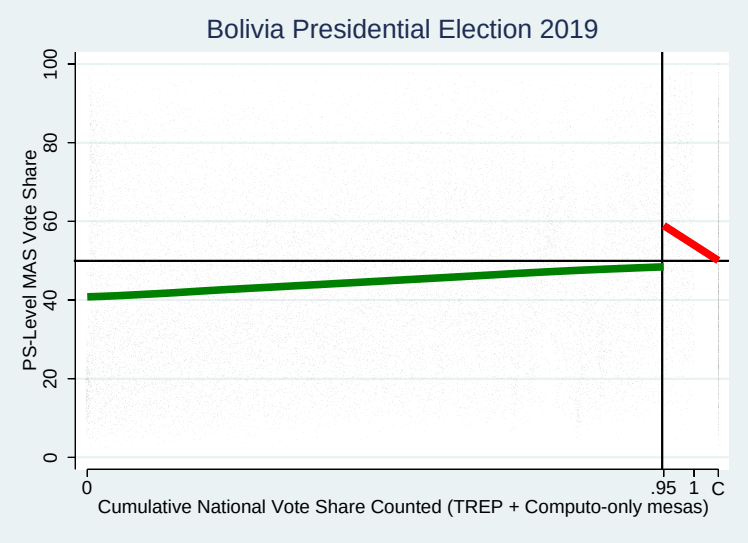
The polling-stations that are only in the Computo data set but not in the TREP do not show as marked an increase, but they are still different than the first 95% of the polling stations included in TREP. But the break between the first 95% of the TREP polling stations and the last 5% of the TREP polling stations is quite apparent.

The figures above relied on the time-stamps reported in the TREP data set. These 1,511 polling stations do not have a time-stamp in the TREP data set. So how could they be included?

For the sake of the exercise, one way could be to append them all to the very end of the count. There are a total of 34,555 polling stations in the TSE data. The 1,511 polling stations that were not in the TREP are 4.3% of that total. The x-axis in the graphs above are based on the cumulative TREP vote count, which runs from 0 (no votes counted) to 1 (all votes counted in TREP). Let's give the 1,511 "missing" polling stations a value of 1.04 (since they are 4% of the cases) for that variable. This is effectively what Idrobo et al. do too.

[1] In the first table on page 90, I pool the last 5% of TREP and the Computo-only polling stations. The column label says this explicitly but perhaps it wasn't clear. But the second table on page 90, reproduced here, makes that unequivocally clear that the 1,511 polling stations that are only in the Computo data set were not excluded from the analysis.

Using a local linear regression and including the Computo-only polling stations, both of which points are the core critiques of Idrobo et al., yields the following graph:



Two points are apparent:

1) The break identified in the original analysis is still apparent at the 95% of the TREP-only cumulative count (0-1 on the x-axis)
2) The addition of 1,511 polling stations at the arbitrary 1.04 (labeled "C" for Computo-only in the graph above) pulls the post-break line down because, as already reported in the second table on page 90 (reproduced above) the Computo-only polling stations, do not show as big a break as the last 5% of the TREP only count.

So, why do Idrobo et al. find something different? I do not know for sure, but I speculate it is because they do the following:

1) Add in the additional 1,511 Computo-only polling stations at the end of the TREP data set by giving it some arbitrary value as I did.
2) Estimate a break at the 95% mark of this new data set.

But the second step would be a mistake.

Why? Recall that the analysis above had identified a possible break point at 95% of the TREP data set, which was 33,044 out of the 34,511 total polling stations. So adding in the additional 1,511 polling stations shifts where the break point falls in the cumulative count. The TREP contains 33044 / 34511 polling stations, or 95.7% of the total polling stations. Therefore, the correct placement of the break once we arbitrarily include the missing polling stations is at 95% of 95.7% (the TREP share of the total distribution), which is **91%** of the TREP+Computo-only (i.e., total distribution).

COMPUTO DATA SET

The TSE also issued a second data set which includes all 34,511 polling stations. The time stamps here are different from the TREP data set. The TREP data set records when the *acta* was received by the TSE; the Computo data set time-stamps records when the *acta* was verified and accepted by TSE as official. These are not the same, though they are loosely correlated as *actas* that were recorded in TREP later are more likely to be confirmed in Computo later.

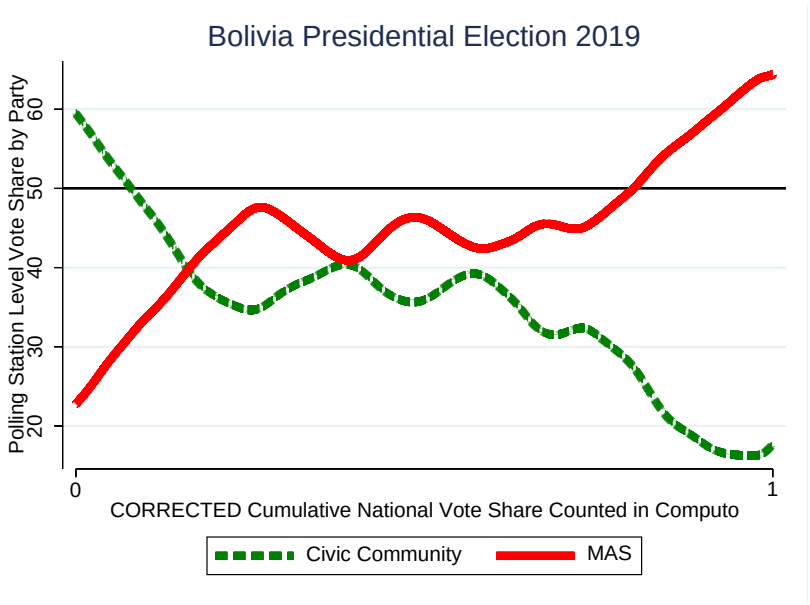
Erratum

In working with the Computo data set, I made a mistake. The Computo time-stamp I used (*ComputoDate*) was in alpha-numeric format, which resulted in an incorrect sorting. This error was identified by David Rosnick of CEPR and publicized on 24 August 2020. I am grateful for the correction.

I have corrected the mistake and re-done the analysis reported originally on pages 92-93 of the OAS report which were based on the Computo data set. As is clear below, none of the results change, and neither do any of the substantive conclusions.

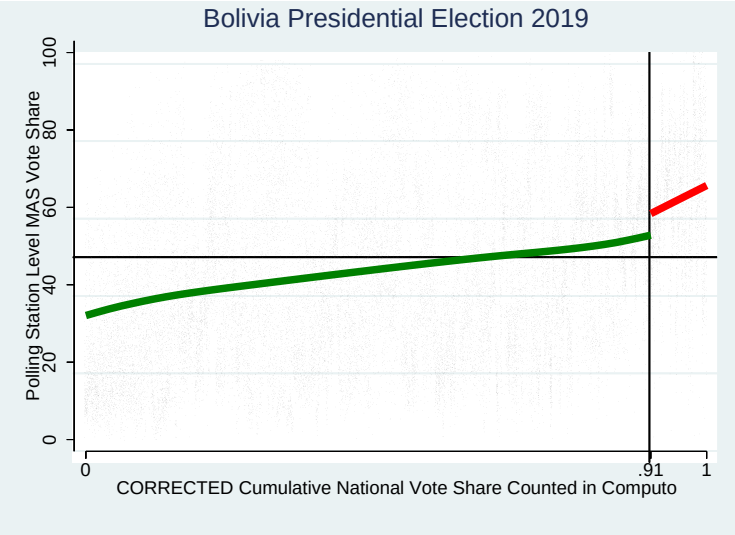
Working with the Corrected Computo Time Stamp

I begin by re-doing the figure originally on page 91 of the OAS report but with the corrected Computo time stamp (*NewComputoDate*).



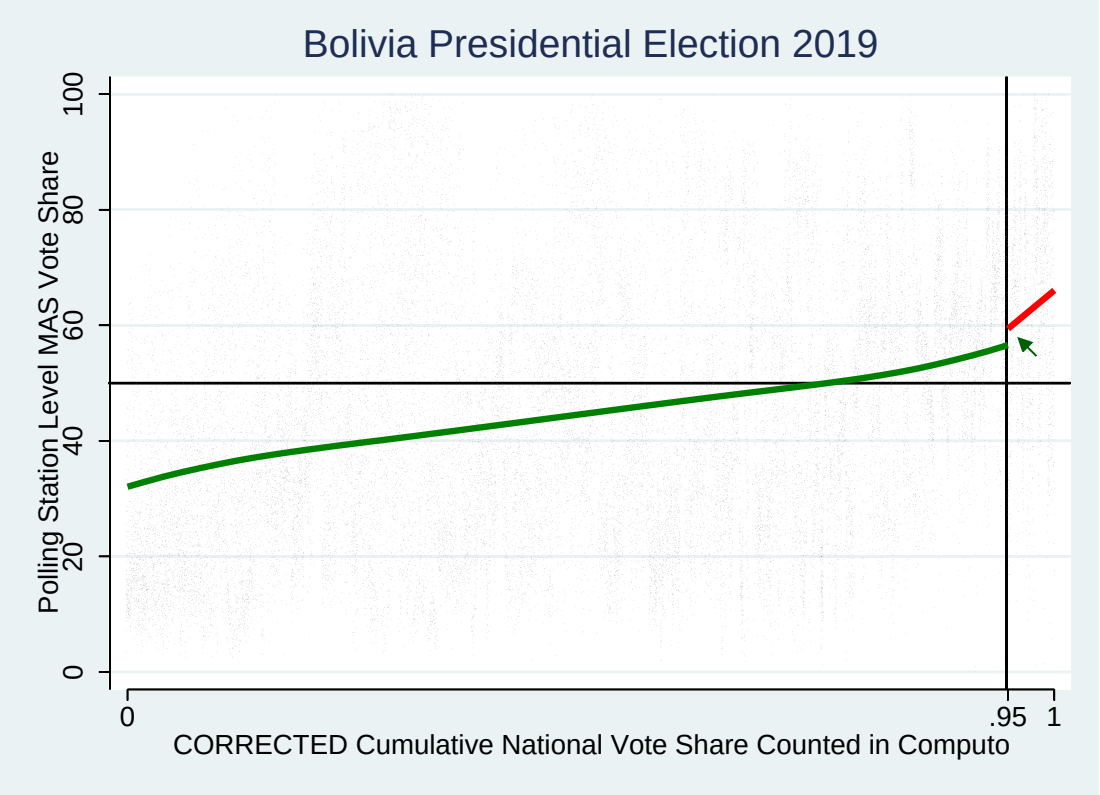
The figure is effectively unchanged from the original on page 91 of the OAS report.

As pointed out above, when working with the full set of 34,555 polling stations, the breakpoint identified in the initial analysis of the TREP-only data set falls at the 91% mark of the total distribution. Here is the Computo data with corrected Computo time stamps:



Yes, that's a break.

Some will worry that I am moving the goalpost by putting the break at 91% of the total distribution. They might ask that I keep it at 95%. OK.



Still a break.

And here is the table originally at the top of page 93, but re-done using the corrected time stamp:

	PS-Level MAS Vote Share		PS-Level CC Vote Share		MAS Advantage over CC		MAS Advantage over CC	
	0-95	95-100	0-95	95-100	0-95	95-100	0-91	91-100
National	44.3	62.8	35.5	16.9	8.8	45.9	7.2	45.6
Beni	33.6	51.3	32.4	19.7	1.2	31.6	1.1	31.6
Chuquisca	34.1	63.7	46.6	19.0	-12.5	44.7	-12.5	44.7
Cochabamba	55.3	-	31.8	-	23.5	-	23.5	30.6
La Paz	50.1	59.0	29.4	18.8	20.7	40.2	15.8	43.3
Potosi	42.7	71.6	33.2	8.7	9.5	62.9	9.5	63.0
Santa Cruz	33.5	37.2	44.5	35.6	-11.1	1.6	-11.1	1.6
Tarija	38.9	-	38.8	-	0.2	-	0.2	-

Source: Computo data and time stamps used to calculate cumulative vote thresholds; author's calculations corrected on 25 Aug 2020

Whether you place the cut-point at 91% or at 95%, there's no escaping the break. The MAS advantage over the CC at the polling station level went from an average of under 9 points for the first 95% of the total vote distribution to five times as much (45.9% vote *margin*) in the last 5 percent. Without that rapid increase, it is unlikely the MAS candidate for President would have cleared the 10% overall vote margin needed to avoid the second-round run-off election.

Summary

The presence of a break is not in and of itself a problem, though experience examining election returns rarely shows a large jump such as the one documented in the table at the top of this page, especially in the final stages of a vote count. At any rate, I stand by my initial conclusion that it merited further investigation. The OAS conducted this investigation using its field teams. My analysis focused their attention on a specific set of polling stations. What they found there is documented in great detail over the first 85 pages of their report. It is those findings that led to their conclusions. If they hadn't found the irregularities they did, the OAS presumably would have reached a different conclusion. As this memo demonstrates, and the accompanying data and code will make clear, my results hold up even with a corrected time stamp. To the extent that my analysis facilitated the uncovering of irregularities that undermined the integrity of the 2019 Bolivian presidential election, I am proud.

Appendix A

Comment on the 2019 Bolivia Presidential Election and OAS Statistical Analysis
Irfan Nooruddin
19 August 2020

On October 20th, 2019, Bolivia held a national election. As part of its process, the OAS commissioned me to conduct an independent analysis of the electoral data from the Bolivian Supreme Electoral Tribunal (TSE). My analysis suggested that there was a change in the vote trends towards the end of the vote count that was critical for putting the incumbent President Evo Morales of the MAS party over the 10% vote margin threshold required to avoid a run-off election under Bolivia's electoral rules. To be clear, Mr Morales had a clear and insurmountable advantage over his principal opponent, Carlos Mesa of the Civic Community party, but with most of the votes counted the advantage was below the 10% mark. Without the rapid increase in Morales's advantage over Mesa at the very end of the vote count, a run-off would have been required. But the rapid increase votes for MAS in the final stages of the counting put Morales over the threshold.

Others have suggested that the rapid gain in MAS votes ought not to have evoked suspicion and that an extrapolation of the vote trends from earlier in the count would have led to the same end result. This argument was first put forth by researchers at the Center for Economic and Political Research (CEPR) in Washington, D.C., and duplicated by researchers from MIT's Election Data Lab. A second study, by researchers at the University of Pennsylvania and Tulane University, argues that I was mistaken in my conclusions, that I made incorrect assumptions in modeling choices, and that they cannot replicate my findings. A story published by the *New York Times* on June 7, 2020, repeated these claims (link).

I have archived all the data and code required to replicate my results at the Harvard Dataverse (https://doi.org/10.7910/DVN/SGOFSC). All my results hold, including when I use the alternative modeling choices suggested by the UPenn/Tulane critics. The mistake, it appears, is entirely theirs because of an incorrect understanding of the methodological choices I used even though I explained these in great detail on the phone to a member of their research team.

It is regrettable that the *New York Times* saw fit to print claims based on an unpublished draft paper that made elementary errors leading to erroneous conclusions and to needless controversy.

In this memo, I detail my analysis and respond to the critics.

Statement of Main Findings

The official report of the OAS is availableat https://www.oas.org/fpdb/press/Audit-Report-EN-vFINAL.pdf.

The code and data archived in the Harvard Dataverse replicates every figure and number reported.

Since all my claims are already in the public domain, I ask interested readers to read pages 86-93 carefully to understand what I said. I do not repeat them here.

Confirmations of Core Findings

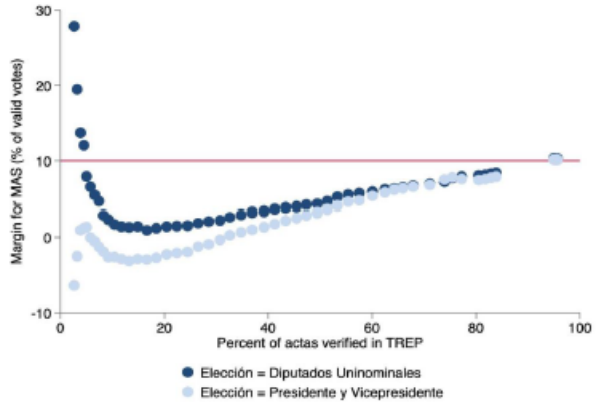
Independent analyses by other researchers reach similar conclusions to mine. I leave it to interested individuals to read their findings and draw their own conclusions. As of this writing, I am aware of the following pieces that corroborate my findings:

1) Diego Escobari and Gary Hoover. "Evo Morales and Electoral Fraud in Bolivia" (Nov 2019)
2) Walter Valdivia and Diego Escobari. "Bolivia's Electoral Fraud Reckoning" (March 2020; $$)
3) John Newman. "The OAS Conclusions about the election integrity of the Bolivian election are correct" (April 2020)

Criticisms of the OAS Audit

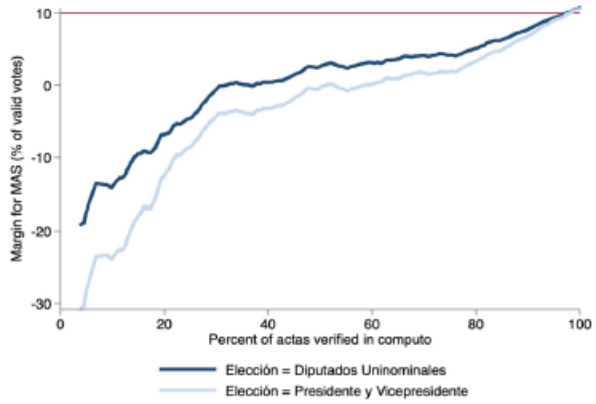
CEPR/MIT Analysis

Researchers from CEPR published a report in November 2019 arguing that the late gains in the MAS vote were consistent with earlier vote trends. Here's the raw figure of the vote count (Figure 1 of their report):



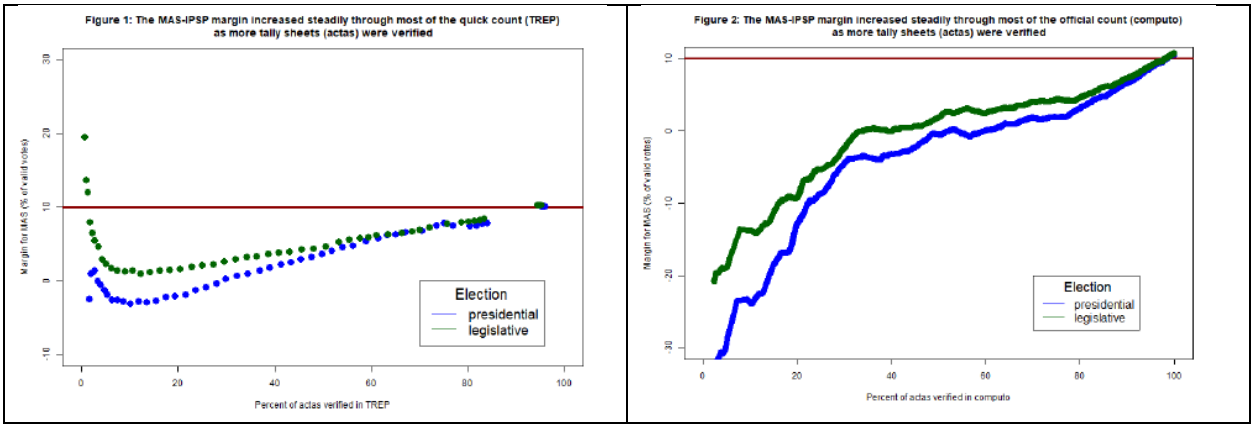
This is based on the TREP data: note the gap in the plot towards the end, starting with approximately 84% of actas verified in TREP. The question is whether it is plausible that Morales got over the 10% threshold.

The core of the CEPR argument is presented in Figure 2 (page 10) of their report in which they use the COMPUTO data:



From my perspective, the slope of the line from approximately 80% to 100% of the cumulative vote count is markedly steeper than over the previous 60% of the cumulative count. In my experience it is rare to see such a steep slope in a cumulative vote count at the tail end of the distribution.

In February 2020, a pair of researchers affiliated with MIT's Election Data and Science Lab published a piece in the Washington Post's Monkey Cage blog stating that they found no evidence of anything problematic either. While a note was later appended to the piece revealing that they had been commissioned by CEPR, readers of their article would have had to click on a link several paragraphs into the piece to find the underlying report on which the Monkey Cage piece was based. Here are Figures 1 and 2 from the Curiel-Williams report:



If they look familiar, it's because they are the same figures as were published three months earlier in the CEPR report. Even the captions are unchanged. Let's just say that if one of my students submitted a paper like this, they'd be facing a honors council hearing.

UPenn/Tulane Analysis

More recently, a separate analysis was made public by researchers affiliated with the University of Pennsylvania and Tulane University. This analysis suggests that my analysis was incorrect because, they allege, I

- Exclude observations in the dataset that accounts for 4.1% of the vote-share; and
- Use an estimator that is unsuited to carry out a regression discontinuity analysis.

When corrected, they claim, my finding of a jump in the MAS vote count is eliminated.

They are wrong on both scores. My most charitable explanation for why they got it wrong is that they did not read carefully enough, and did not understand the data with which they were working.

1) **No observations were excluded.**

The 4.1% of observations they think I excluded are in the dataset and are included in the analysis. How they miss this is beyond me.

There are two data sets relevant here, referred to as TREP and COMPUTO. Both are official data sets published by Bolivia's Supreme Electoral Tribunal. TREP is the preliminary quick-count data; COMPUTO is the final verified count.

1,511 polling stations (4.1% of the total polling stations) were included in the COMPUTO data set but not in the TREP data set. But all the analyses I conducted include them**.**

The figure on page 88 explicitly uses the TREP time stamps. The 1,511 polling stations that are only in the COMPUTO data set do not have TREP time stamps. One cannot include them in the figure, which is why in the tables on pages 89-91 I break out those polling stations separately, explicitly, and transparently.

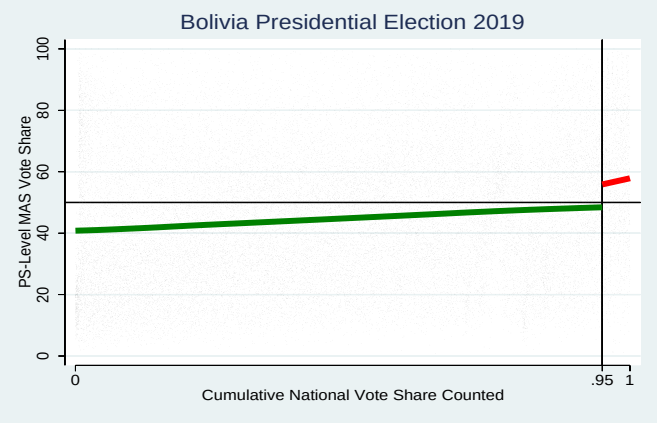
Given that the 1,511 polling stations do not have TREP time stamps, I do not understand how the UPenn/Tulane analysis could possibly include them. The only way to do so would be to place them at the very end of the data set for the purpose of the graph except then we're giving all of these the same arbitrary time count. Instead, one can use the COMPUTO timestamps which exist for all the actas, which, in fact, is what I do on pages 91-92.

In short, this claim by the UPenn/Tulane team is demonstrably incorrect; why they would have made it escapes me.

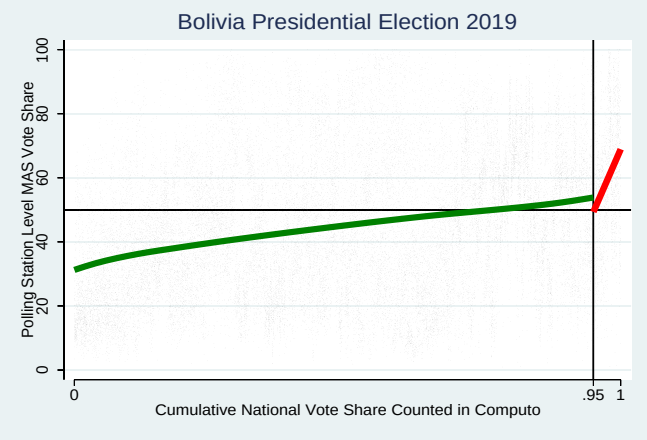
2) Using a different estimator for regression discontinuity doesn't change the finding

The UPenn/Tulane team argue that I use a degree-zero local polynomial estimator (*lpoly*) and that I should have used a linear regression (*lpoly with degree 1*) instead to avoid undesirable properties with degree-zero local polynomials at boundary points. Fair enough, but as I show in the replication materials, using *weighted running lines lowess* (instead of the *running means lowess* that I did use) doesn't alter my findings, and neither does using the linear regression they advise.

Using TREP timestamps with lpoly with degree 1 as requested by UPenn/Tulane researchers



Using COMPUTO timestamps with lpoly with degree 1 as requested by UPenn/Tulane researchers



Conclusions

As a sincere reader of pages 86-93 of the OAS audit report will see, at no point do I ever allege fraud. My findings accompanied 86 other pages of on-ground monitoring by OAS specialists, a fact conveniently ignored by those who would claim that the election had no problems. Academics can and should disagree about modeling choices, but the claims reported in the June NYT article are incorrect and should be withdrawn. Bolivia deserves better.