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It is regrettable that the *New York Times* saw fit to print claims based on an unpublished draft paper that made elementary errors leading to erroneous conclusions and to needless controversy.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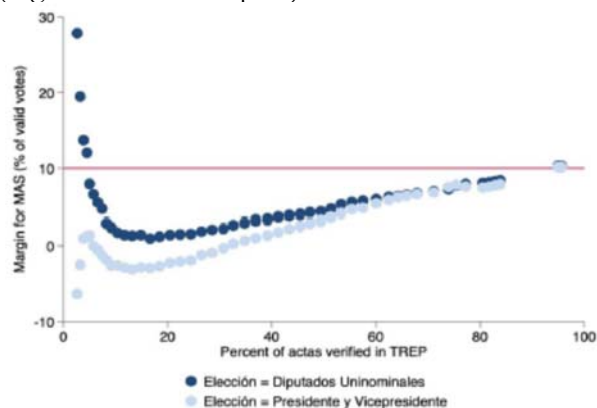

This is based on the TREP data: note the gap in the plot towards the end, starting with approximately 84% of actas verified in TREP. The question is whether it is plausible that Morales got over the 10% threshold.

The core of the CEPR argument is presented in Figure 2 (page 10) of their report in which they use the COMPUTO data:
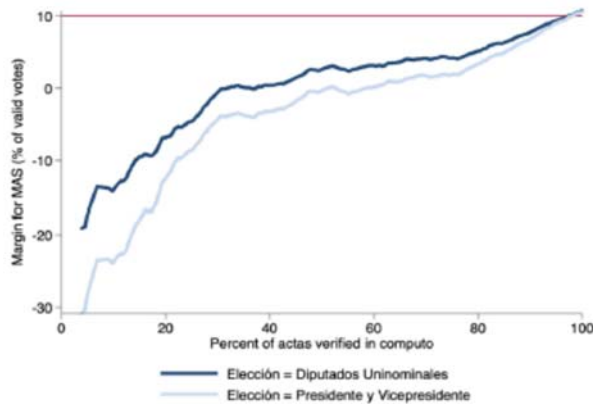
From my perspective, the slope of the line from approximately 80% to 100% of the cumulative vote count is markedly steeper than over the previous 60% of the cumulative count. In my experience it is rare to see such a steep slope in a cumulative vote count at the tail end of the distribution.

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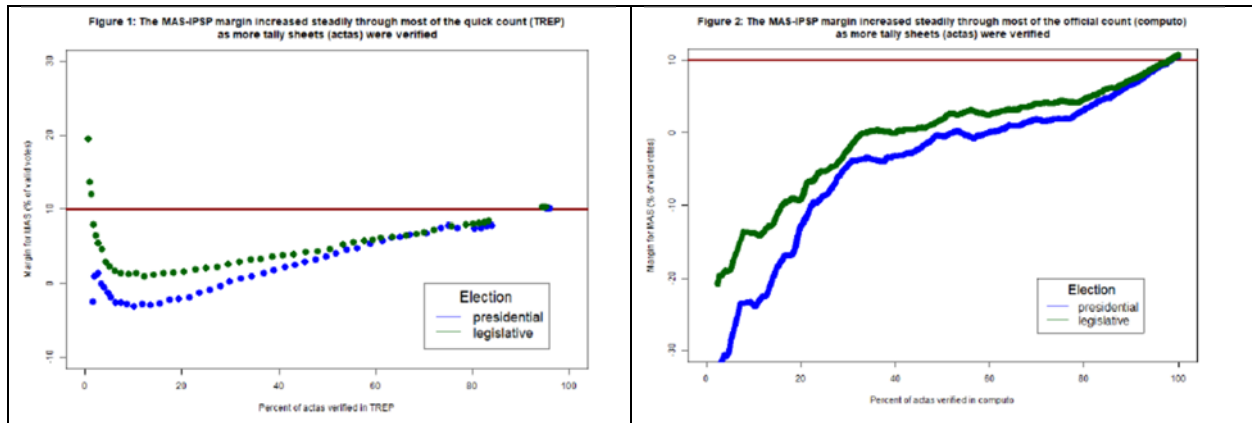
If they look familiar, it's because they are the same figures as were published three months earlier in the CEPR report. Even the captions are unchanged. Let's just say that if one of my students submitted a paper like this, they'd be facing a honors council hearing.

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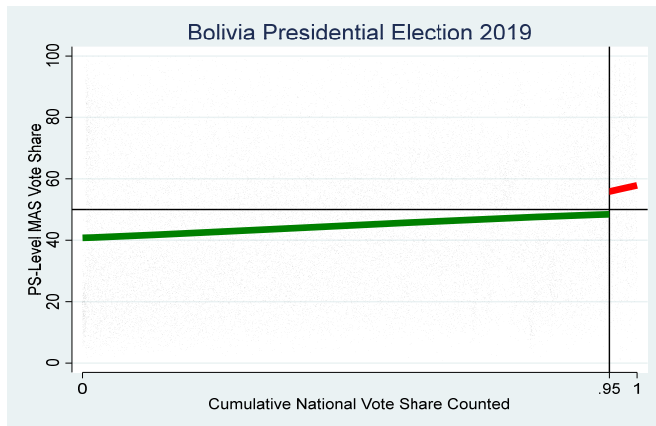
Using COMPUTO timestamps with lpoly with degree 1 as requested by UPenn/Tulane researchers
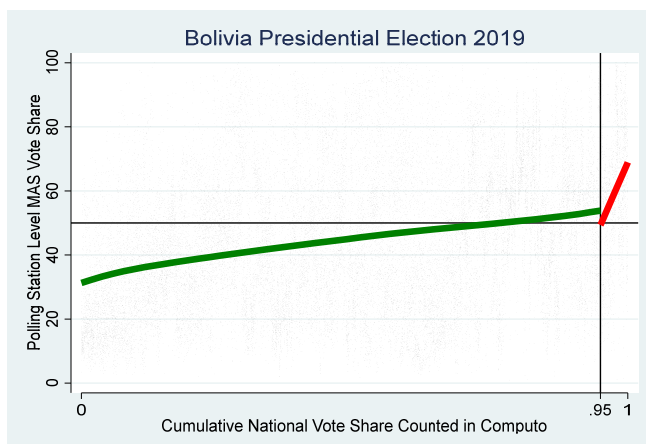
Conclusions

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.