



What Criticisms of Bolivia's 2019 Elections Continue to Get Wrong

By David Rosnick*

November 2020



Center for Economic and Policy Research
1611 Connecticut Ave. NW
Suite 400
Washington, DC 20009

Tel: 202-293-5380
Fax: 202-588-1356
<https://cepr.net>

* David Rosnick is an Economist at the Center for Economic and Policy Research.

Contents

| | |
|--|-----------|
| <u>Introduction</u> | <u>3</u> |
| <u>Newman and a Framework for Analysis</u> | <u>5</u> |
| <u>Inadequate Geographic Controls in Other Analyses</u> | <u>25</u> |
| <u>Analyses of the Official Count</u> | <u>35</u> |
| <u>Further Rebuttal to Newman</u> | <u>43</u> |
| <u>Conclusion</u> | <u>58</u> |
| <u>References</u> | <u>59</u> |
| <u>Appendix: Escobari and Hoover Robustness Check</u> | <u>62</u> |
| <u>Annex: The OAS Statistical Analyses in the Final Report</u> | <u>64</u> |
| Replication | 66 |
| Prediction and Reanalysis | 72 |
| Replication, Prediction, and Reanalysis of Other Figures | 75 |
| Conclusion | 80 |

Acknowledgements

The author wishes to thank Dan Beeton, Karen Conner, Kevin Cashman, Brett Heinz, Jake Johnston, and Guillaume Long.



Introduction

On October 20, 2019, Bolivia held a presidential election. The official count would show incumbent Evo Morales with a first-round victory with 47 percent of the valid vote — more than 10 percentage points ahead of the runner-up, former president Carlos Mesa. However, Morales was not able to finish his term in office — let alone begin his next term. On November 10, under pressure from the military, Morales was forced to resign and flee the country.

Even before the October election, many in the opposition had pledged not to accept the results of the vote if Morales won. The Organization of American States (OAS), in country to observe the vote, poured gasoline on the fire when — the day after the election and with the official count not yet complete — it issued a press release expressing concern about “an inexplicable change in trend [in the preliminary count] that drastically modifies the fate of the election and generates a loss of confidence in the electoral process.”¹ The statement delegitimized the election in the eyes of many, both in Bolivia and internationally. It also provided opposition elements with the justification they needed to reject the election results and force Morales from office three weeks later.²

As CEPR and a number of other independent researchers have shown, the OAS has produced no credible evidence to back up its statement of October 21, 2019. While a long interruption of the preliminary vote count — from around 7:45 p.m. on October 20, 2019 to 6:30 p.m. the following day — raised understandable concerns, there was no “inexplicable” or “drastic” change in trend following the interruption, as the OAS alleged.

Irregularities were going to be inevitable in this election — as with any other — even absent any fraud. However, the OAS’s initial unsubstantiated allegation strongly colored how the OAS interpreted the existence of irregularities, confusion, and destruction of voting materials in the election as evidence of systematic and intentional fraud. The OAS sought out irregularities specifically on tally sheets that heavily favored Morales on the basis of its unjustified claim —

¹ OAS (2019a).

² This is particularly distressing as the preliminary count has, in any case, no official standing. The purpose of this count is to provide the public a day-of sense of how the populace voted. By contrast, the official count is a lengthy process of scrutinizing tally sheets. As a fraud-prevention measure, no corrections are allowed on the tally sheets. Instead, when errors are made, these are presented as “observations” signed and fingerprinted in an allotted section of the sheet. Most frequently, observations report that votes for one party were incorrectly recorded on the tally sheet for another party. These observations are not taken into account in the preliminary count, resulting in inaccurate reporting of results.

and then used the fact of the irregularities to justify its prior actions. Likewise, the OAS held up any sign of confusion or chaos (or even destruction) by any party in the election as further evidence of opportunity for manipulation by the government, seen through the lens of its unjustified claim. The OAS went so far as to bemoan the difficulty in auditing tally sheets that were destroyed when opposition protesters burned several departmental election offices. The audit also included a deeply flawed statistical analysis purporting to substantiate the OAS's initial claim of inexplicable change in the trend of the preliminary vote count. We have reviewed and responded to the audit in detail elsewhere.³

The importance of the OAS's insistence that the late results were inexplicable cannot be overstated. The initial allegation and subsequent audit were cited by the de facto government of Bolivia in their persecution of members of the Tribunal Supremo Electoral (TSE) and prominent members of the political party that Morales led — the Movimiento al Socialismo—Instrumento Político por la Soberanía de los Pueblos (MAS-IPSP, or MAS). Human Rights Watch reported that prosecutors charged at least four members of the Santa Cruz TSE

solely on the report by the Organization of American States ... according to the charging document and the official transcripts of three hearings.⁴

As The New York Times reported, the OAS allegations of fraud “fueled a chain of events that changed the South American nation’s history.”⁵

Since Bolivia’s elections last year, several researchers have performed statistical analyses attempting to show fraud based on a failure to completely explain the “inexplicable change in trend” that so concerned the OAS the day after the vote. Once subject to minimal scrutiny, none of these studies credibly put the official election result in doubt.

In previous papers, we have critiqued the research of Diego Escobari and Gary Hoover, John Newman, and Irfan Nooruddin, the statistician hired by the OAS to perform its own statistical analysis of the vote.⁶ In a document named “Newman Response to Idrobo et al and Rosnick,” Newman responded to our critique of his work.⁷ The OAS has since cited his work as further justification of its actions in Bolivia.⁸

³ Johnston and Rosnick (2020).

⁴ HRW (2020).

⁵ Kurmanaev and Trigo (2020).

⁶ Rosnick (2019) and Rosnick (2020).

⁷ Newman (2020a). His work in question is Newman (2020b).

⁸ OAS (2020a).

In this paper, we provide a detailed rebuttal to Newman’s response. Primarily, we showcase the general inadequacy of his approach and extend his framework for further analysis. With that framework, we revisit and build upon our previous critiques and expand these to include the work of additional researchers, including that of Rómulo Chumacero, and of Edgar Villegas.⁹ In Section I, we highlight a common fault in the primary statistical evidence against Morales. To varying extents, these researchers failed — partially or completely — to account for the tendency of tally sheets representing opposition-heavy precincts to be counted before those in areas supporting Morales. In Section II, we address additional concerns raised by Newman in his response.

In summary, we reaffirm our previous conclusions that Morales’s first-round victory was predictable based on the data available at the time of the October 21, 2019 OAS statement that, in the words of The New York Times, “fueled a chain of events that changed the South American nation’s history.”¹⁰

Newman and a Framework for Analysis

In April, John Newman — former representative of the World Bank in Bolivia — wrote a paper concluding that the OAS was “correct to question the integrity of the 2019 Bolivian General Election.”¹¹ In response, we pointed out two important flaws in his analysis.¹²

First, he employed unnecessary statistical tests of the election data for which the true result was already known. His hypothesis of interest is that the early results (pre-interruption of the preliminary count) differed in distribution from the late results (post-interruption) across opposition-favored geographies. However, by conditioning his sample on early results that favored the opposition, the late results have a relative tendency to favor Morales even in the absence of any fraud. Logically, detecting this tendency in the data does not put the results of the election in doubt.

⁹ Chumacero (2019) and Alvarado, et al. (2019).

¹⁰ Kurmanaev and Trigo (2020).

¹¹ It is not clear if Newman was contracted to perform his analysis, but this paper and Newman’s follow-up appear to be available exclusively on a Google Drive administered by a member of the de facto Bolivian government’s mission to the OAS and promoted on Twitter by the mission. Bolivia en la OEA (2020).

¹² Rosnick (2020).

Second, Newman established a counterfactual estimate of the election outcome, adjusting official results “where there is a suspicion of voter manipulation” inferred illogically on the basis of his tests. He concludes: “The counterfactual estimates suggest that a second round would have had to be held were it not for the change in distributions of margins before and after the cut-off for *municipios*¹³ supporting [Mesa] before the cut-off.” Significantly, the use of *municipios* (municipalities) as his unit of analysis created a bias against Morales as it ignored considerable evidence that Morales’s support came disproportionately late in the count within municipalities — even within localities.¹⁴

Table 1

Administrative Divisions Used in This Report

| Order | Official (TSE) | Label | |
|-------|----------------|-----------------------|------------------|
| | | Spanish | English |
| –1 | | | Foreign/Domestic |
| 0 | Yes | <i>País</i> | Country |
| 1 | Yes | <i>Departamento</i> | Department |
| 2 | Yes | <i>Provincia</i> | Province |
| 3 | Yes | <i>Municipio</i> | Municipality |
| 4 | Yes | <i>Localidad</i> | Locality |
| 5 | Yes | <i>Recinto</i> | Precinct |
| 6 | | <i>Vecino</i> | Adjacency |
| 7 | Yes | <i>Acta (or Mesa)</i> | Tally sheet |

Source: See footnote 13 for details.

In response to our criticism, Newman writes that even if his statistical analysis (the entirety of his initial report) is deficient, the inevitable discovery of irregularities supports his position that the OAS was correct in its snap judgment that the 2019 presidential election quick count in Bolivia was not legitimate.

¹³ The Tribunal Supremo Electoral (TSE) in Bolivia provided hierarchical election results. With increasing order, each official administrative division is embedded completely in the previous one, as in Table 1. Thus, a country consists of one or more departments, a department consists of one or more provinces, and so on. The unofficial division of “adjacency” (*vecino*) refers to any tally sheet along with tally sheets in the precinct numbered one larger or smaller. Improperly, but usefully, in contrast to all other orders of administration, any tally sheet may be part of as many as three *vecinos*. For example, the Cochabamba precinct of Liceo Guindalina Loyoza has eight tally sheets and eight different adjacencies: one consists of tally sheets 30679–80, another of 30679–81, 30680–82, and so forth. This follows Chumacero. Also following Chumacero, we may split administrative divisions at orders 0–4 by dividing the domestic (order –1) division into urban and rural parts. Lower-level divisions do not split, as Chumacero’s definition of urban versus rural is by locality.

¹⁴ Hereafter we will refer to the level of administrative division by its English label, except as included in a quotation.

Newman further reanalyzes the data at the locality level. However, this does not actually address the issue we presented. If, within localities, tally sheets from precincts supporting Morales were counted disproportionately late, then his analysis will underestimate Morales's support and lead to a false conclusion that the first-round victory was illegitimate. By deliberately discarding clear and available information, Newman needlessly produces an inferior estimate which predictably understated Morales's support. In this section, we use Newman's framework as a basis for a more thorough estimate.

Now, suppose that we have lost faith in results counted after the public reporting of results was suspended. Perhaps we suspect that the TSE suspended public reporting in order to give it time to falsify many or all remaining tally sheets in an effort to provide Morales with a first-round victory. If our suspicions are correct, then votes on the tally sheets reported after the interruption should inexplicably favor Morales (or disfavor Mesa) when contrasted against the votes on tally sheets publicly reported prior to the interruption.

Breaking down the vote, we find that 33.7 percent of valid votes were located in precincts where all the votes were counted *before* the interruption (we shall refer to these as *all before*), and another 2.5 percent of valid votes were located in precincts that had not yet reported at all (that is, *all after* the interruption). The remaining 63.8 percent of valid votes were located in precincts that were *split* — partially but not completely counted before the interruption. On average, these *split* precincts were mostly counted early. More than half of all votes that were counted before the interruption correspond to these *split* precincts. Another 13.5 percent of valid votes were counted in these same *split* precincts *after* the interruption, as can be seen in **Table 2**. It is the margins in these *split* precincts with which Newman initially concerns himself.¹⁵

Table 2

Valid Votes by Type of Precinct

| Precinct Type | Valid Votes | Percent of Total |
|----------------|-------------|------------------|
| All Before | 2,067,788 | 33.7 |
| All After | 153,890 | 2.5 |
| (Split) Before | 3,088,170 | 50.3 |
| (Split) After | 827,823 | 13.5 |
| Total | 6,137,671 | 100.0 |

Source: TSE (2020) and author's calculations.

¹⁵ Data used in this report and earlier are available at <https://github.com/ViscidKonrad/Bolivia-Elections-2019>. Real-time election reporting and tally sheet images are available at https://archive.org/details/election_bolivia_2019.

Approximately speaking, if a precinct is *split* — counted partially, but incompletely prior to the interruption — then we would expect each candidate to perform similarly on tally sheets counted late as on those counted before. As the smallest available geographic unit in the election, we expect voters within the same precinct to be similar, socioeconomically, and so have like tendencies to vote in a particular way. Nicolás Idrobo, Dorothy Kronick, and Francisco Rodríguez challenge this assumption, providing evidence that even prior to the interruption, there is a tendency within precincts to increasingly favor Morales over the course of the count.¹⁶ Rómulo Chumacero, likewise, broadly posits that adjacent tally sheets¹⁷ are more similar to each other than to other tally sheets in the same precinct. Regardless, the effect of within-precinct variation is likely small in contrast to the effects discussed below. Ignoring this effect, Chumacero's estimate for Morales's margin is reduced by 0.1 percentage points on tally sheets included in the preliminary count after the interruption. This amounts to a reduction in his estimate of Morales's expected support by 1 switched vote in 14,000 over the entire preliminary count.

As the effect of benign within-precinct variation appears to be neither large nor well-established, we will simplify by assuming explicitly that voters are assigned at random to tally sheets. But for sampling error, then, Morales's support relative to Mesa within a precinct should be the same on tally sheets counted late as on tally sheets counted before.¹⁸

This permits us to form a relatively well-founded expectation of what we might see on a given tally sheet counted after the interruption, provided that the sheet comes from a *split* precinct. We simply draw upon the pre-interruption results in that precinct and extend that to the later tally sheets within the same precinct.

Take, for example, the *split* precinct Colegio Sebastián Pagador in Cochabamba. Prior to the interruption, 1,919 valid votes had been counted there, including 1,205 for Morales and 395 for Mesa. Morales's margin of victory on the early votes at the precinct was 42.2 percent of the valid vote. After the interruption, another 7,861 valid votes were counted. Assuming that

¹⁶ Idrobo, Kronick, and Rodríguez (2020) hypothesize that as voters are assigned to a given tally sheet based on surname, it may be that surname is associated with socioeconomic status within a precinct, correlating both with support for Morales and possibly the likelihood of error on the tally sheet. Tally sheets with errors or observations were set aside for special scrutiny, and therefore delayed in their inclusion in the count. Thus, within a precinct, tally sheets representing voters of lower socioeconomic status may more heavily favor Morales and also be counted later than other tally sheets.

¹⁷ See footnote 16 for an explanation of “adjacent” in this context.

¹⁸ If the effect is real, and late tally sheets within precincts tend to more heavily support Morales, then our simplifying assumption will cause us to underestimate Morales's expected support. In light of this, our assumption is conservative: we are less likely to successfully anticipate Morales's first-round victory.

the late votes were similar to the early ones, we would expect Morales to have received 3,318 more votes than Mesa on the 7,861 votes counted late.

Table 3

| Election Results and Margin Calculation for the <i>Split</i> Precinct Colegio Sebastián Pagador | | | | | |
|---|----------------|-------------|-----------------------------|--------------------|----------------------------|
| | Morales (a) | Mesa (b) | Difference (c) = (a)–(b) | Valid Votes (d) | Margin (e) = 100(c)/(d) |
| (Split) Before | 1,205 | 395 | 810 | 1,919 | 42.21 |
| (Split) After | 4,941 | 1,682 | 3,259 | 7,861 | 41.46 |

Source: TSE (2020) and author's calculations.

In fact, as can be seen in **Table 3**, Morales underperformed a little relative to this expectation, netting only 3,259 more votes than Mesa — or 41.5 percent of the valid votes counted after the interruption.

Consistent with our hypothetical suspicion that the late votes were altered in some way, we could simply try to adjust all the late votes to conform to our expectations based on the early votes. In the case of Colegio Sebastián Pagador, this means adding 59 net votes for Morales. In other *split* precincts, we will subtract votes. In precincts that were counted *all after* the interruption, we have no early 2019 data on which to anchor our expectations, so we will leave these precincts untouched for the time being.

What we cannot do is simply assume that the aggregate late vote over the entire election is similar to the aggregate early vote over the entire election. It is not even correct to assume that the aggregate late vote over all *split* precincts is similar to the aggregate early vote over all *split* precincts. This is because there was an indisputable tendency for tally sheets counted pre-interruption to disproportionately favor the opposition.

While fewer than 20 percent of the valid votes from Morales-heavy Colegio Sebastián Pagador had been counted early, the count of U.A.J.M.S Campus in Tarija was 82 percent complete at the time of the interruption, and these votes broke 40 percentage points in favor of Mesa. Even though the early and late margins within each precinct were similar, their combined late margin (+38) was much larger than their combined early margin (+2), not because of fraud, but simply because the late votes were mostly from Morales-heavy Colegio Sebastián Pagador, as can be seen in **Table 4** below.

Table 4**Early and Late Margins in Two Select Precincts**

| | Valid Votes | | Margin | |
|---------------------------|-------------|-------|--------|--------|
| | Early | Late | Early | Late |
| Colegio Sebastián Pagador | 1,919 | 7,861 | 42.21 | 41.46 |
| U.A.J.M.S. Campus | 1,824 | 398 | -39.64 | -31.12 |
| Total | 3,743 | 8,259 | 2.32 | 37.96 |

Source: TSE (2020) and author's calculations.

In theory, Newman's "Oaxaca-like decomposition"¹⁹ is aimed at quantifying and accounting for this bias respecting the timing of the count. We show here that Newman fails in practice to account for much of this pattern and therefore underestimates Morales's support in the votes counted after the interruption.

Applied at the precinct level, Newman's strategy for detecting potential fraud in the election is to break down quantitatively the overall early and late margins into two parts. The first is the contribution from the shift in cross-precinct weights — that tendency for certain precincts to be counted early, or not. The second is the contribution due to shifts in within-precinct margins — any remaining tendency for the late votes to disproportionately favor one candidate in the latter part of the count after the shift in weights is accounted for.

Such an approach is necessarily limited in that some precincts' votes were counted *all before* or *all after* the interruption. We cannot observe within any of these precincts a shift in margin. Such a decomposition is possible only for *split* precincts — with votes counted both before and after the interruption.

Still, for these split precincts, we may remove from the overall change in margin that portion that we cannot explain by shifting cross-precinct weights. This produces a counterfactual election result. In practice, this circuitous approach of decomposing and reconstituting contributions is not necessary. We may directly substitute where possible the early precinct margin for the actual late margin in the same precinct. This produces directly a numerically identical precinct-adjusted election result.

¹⁹ Newman's "Oaxaca-like decomposition" quantifies the relative contribution to the change in margin between early and late votes attributable to a shift in the mix of geographies. In the case of a precinct-level decomposition, nearly all the change in margin is due to the mix of precincts (tally sheets from opposition precincts generally counted early, and those from Morales precincts disproportionately late). "Oaxaca" here refers to Ronald Oaxaca's 1973 dissertation, but Newman's decomposition should more accurately reference Evelyn Kitagawa (1955), whose writing on the subject in the *Journal of the American Statistical Association* predates that of both Oaxaca and Alan Blinder by 18 years.

Such an analysis for the two precincts discussed above can be seen in **Table 5**.

Table 5

| Early and Late Margins in Two Select Precincts (With Adjustment of Late Margins) | | | | | |
|--|-------------|-------|--------|--------|-----------------|
| | Valid Votes | | Margin | | |
| | Early | Late | Early | Late | Late (Adjusted) |
| Colegio Sebastián Pagador | 1,919 | 7,861 | 42.21 | 41.46 | 42.21 |
| U.A.J.M.S. Campus | 1,824 | 398 | -39.64 | -31.12 | -39.64 |
| Total | 3,743 | 8,259 | 2.32 | 37.96 | 38.27 |

Source: TSE (2020) and author's calculations.

Here, we see that although the late votes broke far more heavily for Morales, his late vote slightly underperformed expectations once the source of the votes is taken into account (38.27 percent compared to an actual 37.96).

In other words, though Morales performed much better on tally sheets counted after the interruption, the *pattern* of votes across the individual precincts hardly changed; it is the changing *importance* of each precinct that accounts for the difference.

A similar analysis for all *split* precincts can be seen in **Table 6**. The combined late votes from all *split* precincts disproportionately favored Morales in the official count (a margin of 19.54 percent of the valid vote, compared to 6.82 percent early in the same precincts). However, if we take each *split* precinct and adjust the late votes there so that Morales's margin over Mesa matches the early vote in the same precinct, then we expect — based on the early margins — Morales' margin on the late votes in *split* precincts to be 18.99 percentage points.

Table 6

| Overall Election Results: Official and with Precinct-Adjustment | | | | |
|---|-------------|------------------|----------|----------------------|
| Precinct Type | Valid Votes | Percent of Total | Margin | |
| | | | Official | Precinct* - Adjusted |
| All Before | 2,067,788 | 33.7 | 9.47 | 9.47 |
| All After | 153,890 | 2.5 | 52.12 | 52.12 |
| (Split) Before | 3,088,170 | 50.3 | 6.82 | 6.82 |
| (Split) After | 827,823 | 13.5 | 19.54 | 18.99* |
| Total | 6,137,671 | 100.0 | 10.56 | 10.49 |

Source and notes: TSE (2020) and author's calculations.

*: data adjusted by precinct.

Thus, the increase in margin across the *split* precincts from 6.82 to 19.54 is almost entirely explained by the fact that votes in Morales-heavy precincts were disproportionately counted

after the interruption.²⁰ Based on the early results, our expectation is that Morales should only have performed slightly below the official result. Overall, this adjustment leads us to expect Morales to win by 10.49 percentage points — only 0.07 percentage points below the official result.

While such a simple approach does leave some votes unexplained, this amount is far too small to put the outcome of the election in doubt. Furthermore, this is exactly the conclusion to which Newman's analysis was leading before he abruptly shifted to an analysis based on wider, more internally diverse geographies.

This shift in focus is very strange. Newman says he does this to “obtain a more complete” picture because the split precincts “constituted only 63.82 percent of the total votes.” This of course exaggerates the scope of the problem, as the precincts counted *all before* account for nearly all of the remaining votes. Only the 2.5 percent of votes coming from precincts counted *all after* are not yet adjusted to any expectations based on the early results. In other words, we already have a nearly complete picture.

In order to put the election result in doubt, our expectations for these last 154,000 votes would have to be that Morales won there by less than 32.57 percentage points, rather than by the 52.12 percentage point margin in the official results. Is such a large difference — 30 times as large as the precinct adjustment in the split precincts — plausible? We cannot rely on early results at the precinct level, as these precincts have no early results. However, some of them have early locality results.

More than 85 percent of the votes in precincts that were counted *all after* the interruption were in localities counted *all after* as well. That is, all the votes in the precincts in these localities were counted *all after*; not a single vote from any of these localities had been counted prior to the interruption.

This leaves a bit more than 22,000 votes in the *all after* precincts for which their locality provided some early results as a basis for comparison. Just as we used the early results in *split* precincts to form expectations about the late results in those precincts, we can leverage the *split* localities to form expectations about the 22,000 late votes there. This adjustment does

²⁰ Thus, the increase in margin across the *split* precincts from 6.82 to 19.54 is almost entirely explained by the fact that votes in Morales-heavy precincts were disproportionately counted after the interruption. Based on the early results, our expectation is that Morales should only have performed slightly below the official result. Overall, this adjustment leads us to expect Morales to win by 10.49 percentage points — only 0.07 percentage points below the official result.

not get us anywhere near what is required to put the election result in doubt, as can be seen in **Table 7**.

Table 7

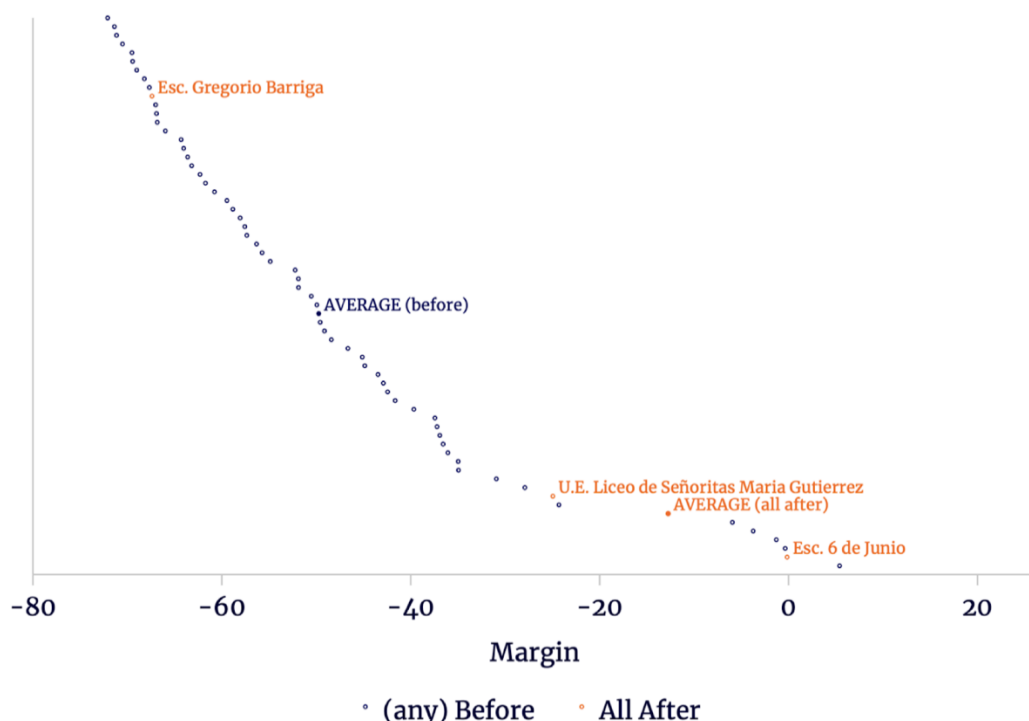
| Locality Adjustment for Precincts that Were Counted <i>All After</i> the Interruption | | | | | |
|---|---------------|-------------|------------------|----------|---------------------------------|
| Precinct Type | Locality Type | Valid Votes | Percent of Total | Margin | |
| | | | | Official | Locality [†] -Adjusted |
| All After | All After | 131,480 | 85.4 | 59.94 | 59.94 |
| | (Split) After | 22,410 | 14.6 | 14.6 | -5.67 [†] |
| Total | | 153,890 | 100.0 | 52.12 | 50.39 |

Source: TSE (2020) and author's calculations.

†: data adjusted by locality.

Note that by nearly 12 percentage points the locality-adjusted margins underestimate Morales's support reported on tally sheets in *split* localities counted after the interruption. By comparison, the underestimate on tally sheets in *split* precincts counted after the interruption was only about half a percentage point. Where available early precinct data is available, it does a much better job of predicting results than where data is only available at the locality. There is good reason to believe this is due to precincts counted *all after* the interruption being fundamentally more favorable to Morales rather than due to a fraudulent exaggeration of support.

For example, consider the split locality of Potosí. On the tally sheets counted prior to the interruption, Mesa won by nearly 50 percentage points. In the precincts counted *all after*, however, Mesa's margin was 37 points lower. The first thing to note is that these late-counted precincts are not especially unusual. Mesa's support in Potosí varied considerably by precinct, even prior to the interruption. Of the 63 precincts shown in **Figure 1**, only three were counted *all after* the interruption. Of the three, one showed relatively low support for Morales and two showed relatively high support for Morales.

FIGURE 1**Results by Precinct (Early and All After) in Potosí Locality**

Source: TSE (2020) and author's calculations.

A locality adjustment, based on the early votes in Potosí, suggests revising down the margins on the three precincts by 37 percentage points. This assumes that the precincts were actually typical for Potosí and not generally more friendly to Morales. However, we have good reason to suspect the reverse. Consider three precincts in Potosí: Esc. Gregorio Barriga, Esc. 6 de Junio, and Esc. Jose A. Zampa. While the correspondence is not perfect, the data clearly suggest that voters in Esc. 6 de Junio are less hostile to Morales than those in Esc. Jose A. Zampa (which voted similarly to Potosí on pre-interruption tally sheets in 2019).

Table 8**Results in Potosí: Precincts Counted All After the Interruption**

| Precinct | Type | Margin (2019) | % Sí (2016) |
|-----------------------|------------|---------------|-------------|
| Esc. Gregorio Barriga | All After | -67.39 | 9.23 |
| Potosí (average) | Before | -49.74 | |
| Esc. Jose A. Zampa | All Before | -49.94 | 12.18 |
| Esc. 6 de Junio | All After | -0.15 | 22.58 |

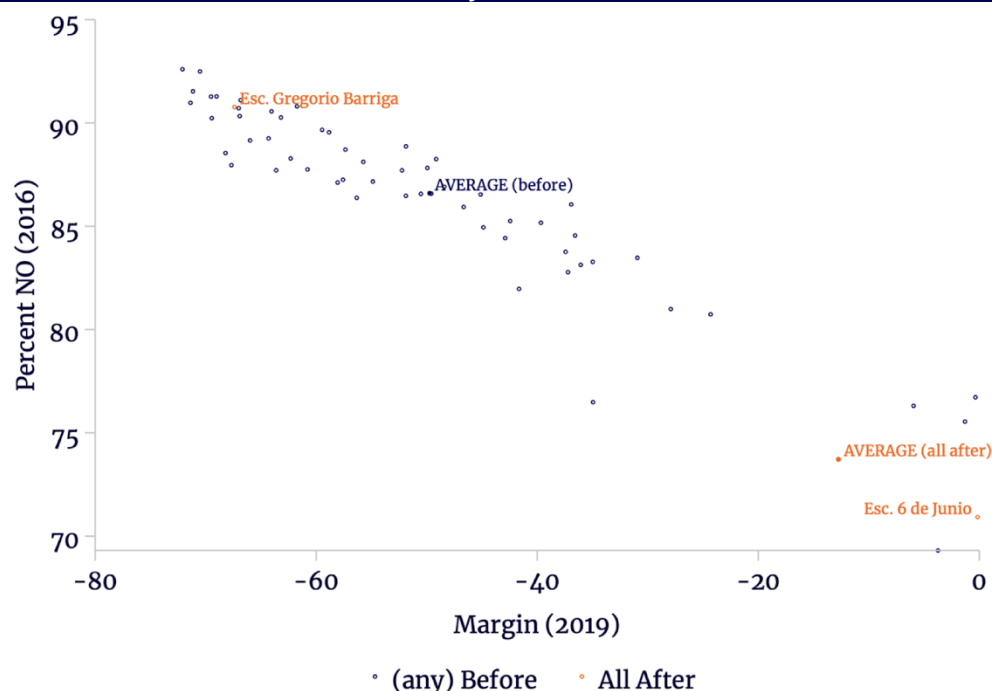
Source: TSE (2020) and author's calculations.

Likewise, Esc. Gregorio Barriga voted more heavily against Morales in both 2016 and 2019. Importantly, if the three late-counted precincts were on average less hostile to Morales than the early tally sheets in Potosí suggest, then a naive locality adjustment will underestimate

Morales's true support. In **Figure 2**, we extend this table to show all precincts in the location of Potosí for which we could find a match to 2016.

FIGURE 2

Votes in 2016 and 2019: Potosí Locality



Source: TSE (2020) and author's calculations.

Based on the early data alone, every percentage point of “NO” vote in 2016 lowered Morales’s margin in 2019 by 3–4 percentage points. This suggests that Morales underperformed expectations in precincts counted *all after* the interruption; rather than –49.74 percent, Morales’s expected margin based on the 2016 vote is close to 0.²¹

We will not concern ourselves here with a precise estimate of the bias in general. We simply note a locality adjustment almost certainly lowers our estimate of Morales’s support in these late-reporting precincts.

In any case, the number of votes at stake in these precincts is relatively low and has little impact on the overall estimates. As can be seen in **Table 9**, below, even though we adjusted the margin on another 22,410 votes from 6.23 percentage points in favor of Morales to 5.67

²¹ We failed to match U.E. Liceo de Señoritas Maria Gutierrez to the 2016 data and so cannot get a reliable estimate for the group of three precincts counted *all after* the interruption. However, the average “No” vote share in the two precincts is 73.7 percent, which translates to a margin between –13.5 and +5.7.

in favor of Mesa, the overall result is still a Morales win by 10.45 percentage points because these precincts represent only 0.4 percent of total votes. Incorporating the early locality results does not put the official election outcome in doubt.

Table 9

Overall Election Results: Official and with Closest Available Adjustment

| Precinct Type | Locality Type | Valid Votes | Percent of Total | Absolute Margin | | |
|----------------|---------------|-------------|------------------|-----------------|--------------------|--|
| | | | | Official | Precinct*-Adjusted | Precinct* or Locality†-Adjusted ^A |
| All Before | | 2,067,788 | 33.7 | 9.47 | 9.47 | 9.47 |
| All After | All After | 131,480 | 2.1 | 59.94 | 59.94 | 59.94 |
| | (Split) After | 22,410 | 0.4 | 6.23 | 6.23 | -5.67 [†] |
| (Split) Before | | 3,088,170 | 50.3 | 6.82 | 6.82 | 6.82 |
| (Split) After | | 827,823 | 13.5 | 19.54 | 18.99 [*] | 18.99 [*] |
| Total | | 6,137,671 | 100.0 | 10.56 | 10.49 | 10.45 |

Sources and notes: TSE (2020) and author's calculations. Note A: Adjustment based on precinct where early precinct data exists (*split* precincts), and is otherwise based on locality where early locality data exists (*split* localities).

*: data adjusted by precinct.

†: data adjusted by locality.

So adjusting precincts counted *all after* the interruption that had early locality data has little effect on the election, yet Newman argues that a locality-based adjustment leads Morales's estimated margin to fall below 10 percentage points. The reason for this difference is that Newman does not adjust late-counted tally sheets in *split* precincts to the early margin in each precinct; he adjusts them to the early margin in the corresponding locality. In other words, he assumes these 827,823 votes should be representative of their localities *even in precincts for which the pre-interruption results say otherwise*.²²

22 In Table 7 of his response, Newman actually adjusts only localities where Morales lost on the pre-interruption tally sheets, so his estimate is a bit larger than in our Table 10 here. If we do likewise, we get 9.73 percentage points, compared to his reported 9.72. The difference is due to slight differences in data definitions. There are six tally sheets included in the last pre-interruption report that Newman counts as post-interruption. He did not include four of the six (2444, 2497–99) because the TSE did not report their transmission timestamps. One (61373) he did not include because — although reported publicly — the TSE verification time stamp is two seconds later than the time stamp of the report. The last (2433) was not included because at the time of the public report it had reported 0 valid votes. The 133 valid votes approved on this tally sheet were originally registered at 12:05:42, but approved with zero votes at 18:32:26; the original 133 valid votes were restored on October 22 at 09:16:16.

Table 10**Overall Election Results: Official and with Closest Available Adjustment (Newman)**

| Precinct Type | Locality Type | Valid Votes | Percent of Total | Absolute Margin | | |
|----------------|---------------|-------------|------------------|-----------------|---|--|
| | | | | Official | Precinct* or Locality [†] -Adjusted ^A | Locality [†] -Adjusted (Newman) |
| All Before | | 2,067,788 | 33.7 | 9.47 | 9.47 | 9.47 |
| All After | All After | 131,480 | 2.1 | 59.94 | 59.94 | 59.94 |
| | (Split) After | 22,410 | 0.4 | 6.23 | -5.67 [†] | -5.67 [†] |
| (Split) Before | | 3,088,170 | 50.3 | 6.82 | 6.82 | 6.82 |
| (Split) After | | 827,823 | 13.5 | 19.54 | 18.99* | 13.08 [†] |
| Total | | 6,137,671 | 100.0 | 10.56 | 10.45 | 9.65 |

Sources and notes: TSE (2020) and author's calculations. Note A: Adjustment based on precinct where early precinct data exists (*split* precincts), and is otherwise based on locality where early locality data exists (*split* localities).

*: data adjusted by precinct.

†: data adjusted by locality.

Newman's expectation is far less well-founded given that we have seen that these late votes come from disproportionately Morales-heavy precincts.

The change in mix of localities among the *split* precincts explains less than half of the change in official margin (6.82 to 19.54). By contrast, the change in mix of precincts explained more than 95 percent of the change. In other words, variation in support among precincts within localities is important for explaining the election results.

To provide a concrete example of how Newman's approach underestimates Morales's support, we reconsider Colegio Sebastián Pagador alongside the early votes in the locality of Cochabamba.²³

Table 11**Election Results and Margin Calculation for the *Split* Precinct Colegio Sebastián Pagador, with Cochabamba Early Votes**

| | Morales (a) | Mesa (b) | Difference (c) = (a)-(b) | Valid Votes (d) | Margin (e) = 100(c)/(d) |
|---------------------------------------|----------------|-------------|-----------------------------|--------------------|----------------------------|
| (Split) Before | 1,205 | 395 | 810 | 1,919 | 42.21 |
| (Split) After | 4,941 | 1,682 | 3,259 | 7,861 | 41.46 |
| Cochabamba Locality (Split) Before | 131,206 | 173,293 | -42,087 | 336,846 | -12.49 |

Source: TSE (2020) and author's calculations.

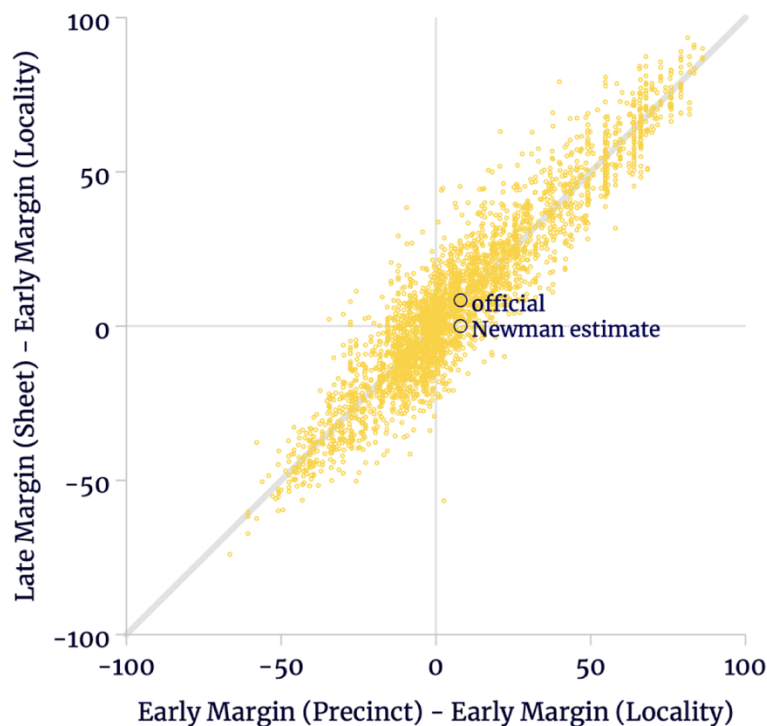
²³ Not to be confused with the municipality, or department, of the same name.

Morales won the early vote at Colegio Sebastián Pagador by 42 percentage points, but lost the early vote in the locality of Cochabamba (encompassing Colegio Sebastián Pagador) by more than 12 percentage points. Given that, should we expect the late results to favor or disfavor Morales? If the answer is not already sufficiently clear that we ought to rely on the early results at the precinct, consider that the precinct voted 64 percent in favor of lifting presidential term limits in the 2016 referendum and 72 percent in favor of Morales in 2014. To suggest that there are literally thousands of *unexpected* late votes for Morales from Colegio Sebastián Pagador alone is, mildly put, not credible.

A broader view of this same effect is visible in **Figure 3**, where *split* precincts are arranged by how favorable or unfavorable the early votes there were relative to the early votes in its corresponding locality. Each point is a tally sheet counted *after* the interruption, showing how favorable or unfavorable the sheet is relative to the early vote in its locality. The early favorability of the precinct is a strong predictor of the favorability of the late tally sheets.

FIGURE 3

Within Localities, Late Results Followed Early Results in Each Precinct



Source: TSE (2020) and author's calculations.

However, we also see that the late tally sheets come from precincts that on average favored Morales pre-interruption by an additional 8 percentage points over their corresponding localities. Newman's locality-level adjustment underestimates Morales's margin there by the

same 8 percentage points.²⁴

There is no defensible reason for ignoring this effect. Newman simply avoids this inevitable conclusion, writing:

If accounting for within-*Localidad* variation would invalidate the [locality-adjusted] results, then eliminating the larger *Localidades* (where the within-*Localidad* variation is likely to be greater than in a smaller, more homogenous *Localidad*) should lead to considerably different results. In fact, it does not.

However, there is no need for an indirect argument. Simply accounting for within-locality variation does invalidate his conclusions. Newman is only correct if Cochabamba (with more than one in nine valid votes counted after the interruption) is not considered a “larger” locality. Even this understates the importance of within-locality variation outside Cochabamba. Cochabamba accounts for only two-thirds of the error from failing to account for the variation among precincts. There are 15 localities with contributions to the error at least as large as that of Trinidad.²⁵ These represent more than 44 percent of the post-interruption votes, and the errors are overwhelmingly one-sided and not infrequently large.

²⁴ The reason we see an eight-point difference here, rather than the six-point difference in Table 10 is that, for clarity, Figure 3 excludes localities with only a single precinct counted early (and therefore have early margins in precinct and locality that are identical by construction).

²⁵ Of those included in Newman’s analysis of select localities (Table 10), Trinidad has the smallest absolute contribution to the error.

Table 12**Contributions to Underestimate of Morales's Support by Locality**

| Locality | Valid Votes (Post Interruption) | | Early Margin | Error in Estimate ^A | Contribution to Overall Error | |
|-------------------|---------------------------------|--------------------|--------------|--------------------------------|-------------------------------|------------------------------|
| | Total | Cumulative Percent | | | Total | Cumulative Percentage Points |
| Cochabamba | 113,716 | 11.6 | -12.49 | 28.76 | 0.53 | 0.53 |
| Tarija | 30,294 | 14.7 | -26.33 | 12.58 | 0.06 | 0.60 |
| El Alto | 65,645 | 21.4 | +31.10 | 4.14 | 0.04 | 0.64 |
| Sucre | 8,453 | 22.2 | -37.80 | 29.69 | 0.04 | 0.68 |
| Potosí | 38,502 | 26.1 | -49.74 | 5.50 | 0.03 | 0.71 |
| Oruro | 14,857 | 27.7 | -3.97 | 10.97 | 0.03 | 0.74 |
| La Paz | 25,729 | 30.3 | -9.71 | 4.66 | 0.02 | 0.76 |
| Quillacollo | 13,053 | 31.6 | +2.70 | 6.44 | 0.01 | 0.77 |
| Santa Cruz | 86,252 | 40.4 | -27.56 | 0.89 | 0.01 | 0.79 |
| Riberalta | 7,915 | 41.2 | -27.13 | 8.69 | 0.01 | 0.79 |
| Quintanilla | 7,279 | 41.9 | -1.41 | 7.52 | 0.01 | 0.81 |
| Yacuiba | 4,797 | 42.4 | -11.14 | 10.37 | 0.01 | 0.82 |
| Trinidad | 7,331 | 43.2 | -6.79 | 6.04 | 0.01 | 0.82 |
| Sacaba | 6,525 | 43.8 | +30.87 | -9.88 | -0.01 | 0.82 |
| Quijarro | 3,141 | 44.2 | -7.40 | -24.59 | -0.01 | 0.80 |
| <i>All others</i> | 548,224 | 100.0 | +38.19 | -0.03 | -0.00 | 0.80 |

Source and notes: TSE (2020) and author's calculations. Note A: Difference between estimate including accounting of precinct variation and estimate without.

On the whole, failure to account for variation among precincts underestimates Morales's election-wide margin by 0.80 percentage points — enough to reach a false conclusion regarding the election. Instead of following through with analysis, Newman assumes an agnostic position:

However, it is also possible that there was a substantial change in margins in Cochabamba and that its experience should not be dropped from the experience of other Localidades in the analysis of potential differences in margins before and after the cut-off.

There is indeed no reason to treat Cochabamba differently than other localities in this respect. The experience of other localities — particularly large localities — is that precinct

variation matters.²⁶ All *split* localities should be treated similarly, with full accounting for within-locality variation. A precinct-level analysis is appropriate for all such localities and not just Cochabamba. It is clear what effect including this information has on the estimates: it utterly undermines Newman's conclusions. We must not project late votes without taking into account patterns across precincts within localities.

Looking back to Table 10, incorporating early locality data to estimate votes in precincts counted *all after* the interruption still leaves 131,480 late votes. These are precincts in localities counted *all after*.²⁷ Of course, localities counted *all after* are made up exclusively of precincts counted *all after*. Some of these localities lie in municipalities counted *all after*, but some will lie in *split* municipalities (with some data counted *before*). This allows us to form expectations about some *all after* localities.

In the same way that we exploited split localities to project some *all after* precincts, we may leverage *split* municipalities to project results in localities counted *all after* the interruption. This still leaves some votes in *all after* municipalities which requires leveraging of early results within provinces, and so on. This is expressly the simple exercise we performed in our report, "What Happened in Bolivia's 2019 Vote Count?" back on November 8, 2019.²⁸

We present a similar analysis in the following tables, working from the largest geographies to the smallest. For consistency with later analysis, we employ Chumacero's urban/rural divide.²⁹

²⁶ It is a nearly trivial exercise — the addition of a single line of code to Newman's analysis — to focus exclusively on the single locality. The counterfactual final margin for Cochabamba is -4.94 percentage points, compared to the official -5.03. This means that once the share of votes coming from each precinct is accounted for, Morales actually underperformed expectations by about 0.4 percentage points.

²⁷ There are a total of 153,890 in "precincts counted *all after*," but using early location data only gets us estimates for the 22,410 (those being split in localities), leaving another 131,480 in *all after* localities.

²⁸ Long, et al. (2019). This simple exercise should not be confused with the more rigorous analysis we described in that report's Data Appendix. The latter provided a range of estimates somewhat less naively constructed. Nevertheless, based on the naive approach, a Morales win by more than the necessary 10 percentage points was expected.

²⁹ The split occurs as a negative-order administrative level distinction. Effectively, Bolivia is split into two different countries, with certain localities defined as part of "urban" Bolivia and others as "rural" Bolivia. Practically speaking, we have doubled the effective number of departments in Bolivia. For example, we distinguish "urban" Beni from "rural" Beni. Likewise, we have in effect doubled the number of provinces and municipalities. At the author's request, Chumacero's urban/rural classification by tally sheet was provided by email.

Table 13**Overall Election Results: Provinces Counted *All After***

| Department Type | Country Type | Valid Votes | Percent of Total | Margin | |
|-----------------|---------------|-------------|------------------|----------|----------------------------------|
| | | | | Official | Department* or Country†-Adjusted |
| All After | All After | 124 | 2.6 | -54.84 | 30.25 |
| | (Split) After | 2,469 | 51.1 | 8.71 | 13.72 [†] |
| (Split) After | | 2,238 | 46.3 | 58.27 | 55.41 [*] |
| Total | | 4,831 | 100.0 | 30.04 | 33.46 |

Source and notes: TSE (2020), Chumacero (2019), and author's calculations. Adjustment based on department where early department data exists (*split* departments), and otherwise based on country where early country data exists (*split* country). For completeness, adjustment in Colombia (counted *all after*) based on the pre-interruption results on tally sheets originating outside Bolivia.³⁰

*: data adjusted by department.

†: data adjusted by country.

We may incorporate these results into our analysis of election results in localities counted *all after*. This expands the analysis to cover 2.1 percent of all votes in the election. Overall, Morales overperformed a little relative to expectations, but again, his very strong performance in these late-counted localities was predictable.

Table 14**Overall Election Results: Localities Counted *All After***

| Municipality Type | Province Type | Valid Votes | Percent of Total | Margin | |
|-------------------|---------------|-------------|------------------|----------|-------------------------------------|
| | | | | Official | Municipality* or Province†-Adjusted |
| All After | All After | 4,831 | 3.7 | 30.04 | 33.46 |
| | (Split) After | 2,104 | 1.6 | 39.02 | 47.78 [†] |
| (Split) After | | 124,545 | 94.7 | 61.46 | 55.72 [*] |
| Total | | 131,480 | 100.0 | 59.94 | 54.77 |

Source and notes: TSE (2020), Chumacero (2019), and author's calculations. Adjustment based on municipality where early municipality data exists (*split* municipalities), and alternatively based on province where early province data exists (*split* provinces). Otherwise, as in Table 13.

*: data adjusted by municipality.

†: data adjusted by province.

Finally, we may incorporate these results into a picture of the overall election, as can be seen in **Table 15**. Though some votes are unexplained in this simple approach, the official results conform closely to expectations, if perhaps a bit less well in the *split* localities.

³⁰ While the early external vote strongly favored Morales, voters in Colombia favored Mesa by a large margin. Thus, the use of early votes from outside Bolivia hugely overestimates the margin in Colombia, by about 87 percentage points. However, the absolute difference in net votes (108) comes to less than 0.002 percent of the total valid votes in the election. This is far too small to make a difference in the analysis.

Table 15

Overall Election Results: Complete

| Precinct Type | Locality Type | Valid Votes | Percent of Total | Margin | |
|----------------|---------------|-------------|------------------|----------|--|
| | | | | Official | Precinct*, Locality†, or Otherwise‡-Adjusted |
| All Before | | 2,067,788 | 33.7 | 9.47 | 9.47 |
| All After | All After | 131,480 | 2.1 | 59.94 | 54.77† |
| | (Split) After | 22,410 | 0.4 | 6.23 | -5.67† |
| (Split) Before | | 3,088,170 | 50.3 | 6.82 | 6.82 |
| (Split) After | | 827,823 | 13.5 | 19.54 | 18.99* |
| Total | | 6,137,671 | 100.0 | 10.56 | 10.34 |

Source and notes: TSE (2020), Chumacero (2019), and author's calculations. Adjustment based on precinct where early precinct data exists (*split* precincts), and alternatively based on locality where early locality data exists (*split* localities). Otherwise, as in Table 14.

*: data adjusted by precinct.

†: data adjusted by locality.

‡: data otherwise adjusted.

At this point, we have an estimate of the full election results. Making no effort to incorporate prior election data, our estimate is closer to the official result than it is to the 10 percentage point threshold. Though limiting our understanding only to the 2019 data underestimates Morales's support, that support is nevertheless consistent with a legitimate first-round victory.

Comparing this result to Newman's locality-level adjustment, we see that within-locality variation accounts for more than 100 percent of the difference. Again, in his original paper, Newman recognized that the shift in precincts explained almost all of the increase in Morales's margin, but deliberately ignored his own finding when constructing his estimates for the election.

Table 16

Overall Election Results: Comparison to Newman

| Precinct Type | Locality Type | Valid Votes | Percent of Total | Margin | |
|----------------|---------------|-------------|------------------|--|--|
| | | | | Precinct*, Locality [†] , or Otherwise [‡] -Adjusted | Locality [†] -Adjusted (Newman) |
| All Before | | 2,067,788 | 33.7 | 9.47 | 9.47 |
| All After | All After | 131,480 | 2.1 | 54.77 [‡] | 59.94 |
| | (Split) After | 22,410 | 0.4 | -5.67 [‡] | -5.67 [‡] |
| (Split) Before | | 3,088,170 | 50.3 | 6.82 | 6.82 |
| (Split) After | | 827,823 | 13.5 | 18.99 [*] | 13.08 [†] |
| Total | | 6,137,671 | 100.0 | 10.34 | 9.65 |

Source and notes: TSE (2020), Chumacero (2019), and author's calculations.

*: data adjusted by precinct.

†: data adjusted by locality.

‡: data otherwise adjusted.

Had the mix of precincts within each locality been unchanged before and after the interruption, Newman's analysis might have been more effective. Larger sample sizes in working with localities might have reduced uncertainty relative to a precinct-level analysis. However, with tally sheets from opposition precincts counted disproportionately prior to the interruption, there is simply no justification for anchoring expectations to the early locality results where early precinct results are available.

Nor does observing that the effect in Cochabamba is strong absolve one of the need to incorporate the precinct-level results. Cochabamba, as an example, clearly demonstrates that precinct-level analysis is critical for avoiding a large and totally unnecessary anti-Morales bias to the numbers.

Inadequate Geographic Controls in Other Analyses

While Newman's approach to evaluating the official election results clearly failed, he was not the first researcher to suggest that Morales's first-round victory was fraudulent while failing to take into consideration the geographic bias in the early count.

No matter the researcher or the specifics of the approach, one finds that when within-locality variation is taken into account, the election results stand up to scrutiny. Only by ignoring this important source of variation are they able to argue that the election outcome was in doubt.

Irfan Nooruddin, on behalf of the OAS, fails to look any deeper than the department level and thereby fails entirely to identify the source of the late break for Morales:

Even if the late-reporting stations [tally sheets] were favorable to Morales and MAS, the size of the advantage MAS enjoyed, in every department, over the late-reported polling stations is extremely unusual.

Nooruddin offers nothing to support his conclusion that there is anything unusual, apart from the fact itself that the late-counted votes favored Morales more than the early. For example, Nooruddin might have tried to understand the late surge in Beni — the first department in his table at the top of page 90 in the OAS report. Looking at Nooruddin's table, he shows that Mesa's lead over Morales over the first 95 percent of the preliminary count was entirely wiped out in the last 5 percent.

Table 17**Beni: Results Across Nooruddin's 95/5 Split**

| Precinct Type | Locality Type | Valid Votes ^A | Percent of Total | Absolute Margin | |
|-----------------|---------------|--------------------------|------------------|-----------------|--------------------|
| | | | | Official | Adjusted |
| All Before | | 79,101 | 36.8 | -4,974 | -4,974 |
| All After | All After | 13,750 | 6.4 | 6,280 | 6,521 [‡] |
| | (Split) After | 898 | 0.4 | 142 | -10 [†] |
| (Split) Before | | 103,396 | 48.2 | -1,585 | -1,585 |
| (Split) After | | 17,554 | 8.2 | 487 | 808 [*] |
| Total | | 214,699 | 100.0 | 350 | 761 |
| Note: Before 95 | | 182,497 | 85.0 | -6,559 | -6,559 |
| Note: After 95 | | 32,202 | 15.0 | 6,909 | 7,320 |

Sources and notes: TSE (2020), Chumacero (2019), and author's calculations. Note A: Valid votes in the OAS table differ slightly because they reflect the valid vote totals as written on the tally sheets rather than the sum of all valid votes cast. The latter is the legal basis for determining the election outcome.

*: data adjusted by precinct.

†: data adjusted by locality.

‡: data otherwise adjusted.

Again, this was a predictable outcome. In **Table 17**, we predict the results on tally sheets not included in Nooruddin's "Before 95%."³¹ For comparison to Nooruddin's table, we report the absolute vote margins (Morales votes, less Mesa votes) rather than the difference as a share of valid votes. We find that in Beni, Morales actually underperformed expectations by 411 votes once geographic differences within Beni are taken into account. Again, we don't take this as an indicator of possible fraud against Morales, but it certainly shows that there is no unexpectedly large swing toward Morales on the late tally sheets from Beni.

Key to understanding this shift is the considerable shift in counting from urban to rural votes within Beni. All 13,750 votes in localities counted *all after* the 95 percent mark came from rural Beni. Forty-eight percent of votes counted in Nooruddin's last 5 percent were from rural Beni, where Morales won early by 37 percentage points. In the first 95 percent, the votes were 93 percent urban; Morales lost these early urban voters by 7 percentage points.

By failing to explain this overall shift in support — merely pointing it out and implying that the shift is suspicious — Nooruddin misleads.

³¹ Nooruddin's "Before 95%" do not quite correspond to the first 95 percent of results made available in the public data over the progression of the quick count. Nooruddin includes tally sheets in the count as they are initially verified. However, the authorities delayed many of these verified results from entry into the public data until they could be scrutinized. Most of these tally sheets were not approved until the very end of the count.

Extending this analysis to all departments and the external vote, we find that 98 percent of all valid votes were on tally sheets counted either before Nooruddin's 95 percent mark or in precincts that had been partially counted at that point. Nearly 80 percent of all valid votes in the "After 95" group were predictable based on early precinct data. This means we should be able to establish well-founded expectations of how the election would proceed past Nooruddin's 95 percent mark.

Table 18

| Overall Election: Results Across Nooruddin's 95/5 Split | | | | | |
|---|---------------|--------------------------|------------------|-----------------|---------------------|
| Precinct Type | Locality Type | Valid Votes ^A | Percent of Total | Absolute Margin | |
| | | | | Official | Adjusted |
| All Before | | 3,280,467 | 53.4 | 267,343 | 267,343 |
| All After | All After | 102,864 | 1.7 | 62,071 | 57,821 [†] |
| | (Split) After | 10,872 | 0.2 | 469 | 1,234 [†] |
| (Split) Before | | 2,318,565 | 37.8 | 222,587 | 222,587 |
| (Split) After | | 424,903 | 6.9 | 95,969 | 94,301 [*] |
| Total | | 6,137,671 | 100.0 | 648,439 | 643,286 |
| Note: Before 95 | | 5,599,032 | 91.2 | 489,930 | 489,930 |
| Note: After 95 | | 538,639 | 8.8 | 158,509 | 153,356 |

Sources and notes: TSE (2020), Chumacero (2019), and author's calculations. Note that "Before 95" contains fewer than 95 percent of valid votes. This is because the 95 percent mark chosen by Nooruddin is of the votes verified in the preliminary count, which did not cover the entire election. Tally sheets not in the preliminary count ("not in TREP") are included in the "After 95." Note A: Valid votes in the OAS table differ slightly because they reflect the valid vote totals as written on the tally sheets rather than the sum of all valid votes cast. The latter is the legal basis for determining the election outcome.

*: data adjusted by precinct.

†: data adjusted by locality.

‡: data otherwise adjusted.

Morales did wind up overperforming expectations, but not by much. With only an 8.75 percentage point lead on the first 95 percent, Morales led by 29.42 percentage points on the rest instead of an expected 28.47. In Newman's terms, 95 percent of the change in Morales's margin is easily explained by geographic differences with no recourse to considering previous elections.

In total, we expect in this simple exercise for Morales to have increased his margin from 8.75 percentage points to 10.49 — more than enough for a first-round victory. Nooruddin's claim that the late break "is extremely unusual" is belied by the data showing clearly that a large swing was predictable.³²

³² In the Annex, we present a more thorough examination of Nooruddin's statistical analyses on behalf of the OAS.

While Nooruddin inexplicably ignored important information within departments, provinces, municipalities, and localities, at least he recognized the importance of differences between departments. Escobari and Hoover, meanwhile, explicitly contradict their own results regarding the importance of any such variation, concluding:

We obtain a point estimate that shows that fraud increased the gap between MAS and CC by *at least 2.67% of the valid votes*.³³ [*Emphasis added.*]

First, they find no fraud whatsoever. Their model permits only one possible explanation for the shift in support: the tally sheets were counted late. Thus, Escobari and Hoover explicitly conclude there was significant fraud because the late tally sheets more heavily favored Morales. This is not analysis; it is begging the question.

For our purposes here, we note that they characterize this estimate as a lower bound on the extent of fraud, but once they repeat the analysis using precinct-level controls, they find that Morales's first-round victory was predictable. Their own analysis shows that geographic factors alone explain almost all of the 2.67 percent difference.

Though Escobari and Hoover present regression models, we may summarize their results within Newman's framework. Simply, as the geographic units under observation get smaller, more and more of the late swing in support is explained.

Table 19

Overall Election: Escobari and Hoover

| Precinct Type | Valid Votes | Percent of Total | Escobari and Hoover Estimates by Level of Analysis | | | Official Margin |
|---|-------------|------------------|--|--------------|----------|-----------------|
| | | | None (Before/ After Only) | Municipality | Precinct | |
| All Before | 2,067,788 | 33.7 | 9.47 | 9.47 | 9.47 | 9.47 |
| All After | 153,890 | 2.5 | 7.88 | 36.75 | 52.12 | 52.12 |
| (Split) Before | 3,088,170 | 50.3 | 6.82 | 6.82 | 6.82 | 6.82 |
| (Split) After | 827,823 | 13.5 | 7.88 | 13.47 | 18.99 | 19.54 |
| Total | 6,137,671 | 100.0 | 7.88 | 9.36 | 10.49 | 10.56 |
| <i>Underestimate^A (% of after votes)</i> | | | 16.77 | 7.53 | 0.47 | |

Sources and notes: TSE (2020) and author's calculations. Note A: Escobari and Hoover (2019; Table 3, Columns 1–3) differ slightly because they do not weigh tally sheets by the number of valid votes and denominate by the valid votes listed on tally sheets, rather than by the legally binding sum of valid votes counted. This has implications for the statistical significance of their findings (see Appendix).

³³ Escobari and Hoover (2019).

Again, Escobari and Hoover’s analyses confirm that the *pattern* of votes across precincts observed prior to the interruption carries over into the post-interruption results. It is only the shift in *importance* — which precincts get counted when — that is unexplained. By labeling this shift “fraud,” Escobari and Hoover mislead. Analyzing at the precinct level gives a more complete picture of the election, and Morales’s first-round victory is predictable.

Edgar Villegas Alvarado³⁴ likewise performs a partial analysis and argues that though Morales’s margin ought to have grown post-interruption, the late ballots ought not to have sufficed for a first-round victory. However, this is again due to a failure to account for all the available information.

The Villegas analysis divides tally sheets into only three geographical categories: the nine department capital municipalities and El Alto (“Ciudad”), the rest of Bolivia (“Provincia”), and votes coming from other countries (“Extranjero”). While rural and urban have been employed as shorthand locating support for Morales and the opposition, respectively, this was never precise. Morales and MAS have always been popular in large parts of the capital municipality of Cochabamba as well as in El Alto. While it may make sense to lump together the capitals and El Alto for purposes of determining urban versus rural geography, it is plainly wrong as a proxy for support for different candidates.

Figure 45 of the Villegas report hints at a compounding error. That figure shows the proportion of tally sheets coming from each of the three geographies both before and after the interruption. Unfortunately, these proportions are used as uniform weights for projecting the late vote; the proportion of valid votes — not tally sheets — in each geography is appropriate. Villegas also uses the “Votos Válidos” check on the tally sheet as the denominator for vote shares and margins where the actual sum of valid votes is the appropriate, legal, basis for computing shares and margins. Correcting for both of these oversights yields the following table.

³⁴ The report attributes this section (16) to an anonymous “postdoctoral investigator in the field of Biotechnology at an internationally renowned University.” For simplicity, we will assign authorship to Villegas: Alvarado, et al. (2019).

Table 20

| Proportions in Villegas's Three Geographies | | | | |
|---|----------------------------|-----------|-------------------|-----------|
| | Pre-Interruption ("TREPo") | | Post-Interruption | |
| | Villegas Fig. 45 | Corrected | Villegas Fig. 45 | Corrected |
| Extranjero | 4.46 | 3.29 | 3.87 | 3.28 |
| Provincia | 42.70 | 40.06 | 58.57 | 55.66 |
| Ciudad | 52.85 | 56.65 | 37.56 | 41.06 |

Sources: TSE (2020) and author's calculations.

Villegas underweights the urban vote, biasing results in favor of Morales, yet still underestimates Morales's support post-interruption because, again, there is considerable variation within these geographies. In **Table 21**, we break down the Villegas analysis by type of precinct and locality as before.

Table 21

| Overall Election Results: Including Full Adjustment and Villegas's Estimate | | | | | | |
|---|---------------|-------------|------------------|-----------------|--|--------------------|
| Precinct Type | Locality Type | Valid Votes | Percent of Total | Absolute Margin | | |
| | | | | Official | Precinct*, Locality†, or Otherwise‡-Adjusted | Villegas§ Estimate |
| All Before | | 2,067,788 | 33.7 | 9.47 | 9.47 | 9.47 |
| All After | All After | 131,480 | 2.1 | 59.94 | 54.77 [†] | 30.17 [§] |
| | (Split) After | 22,410 | 0.4 | 6.23 | -5.67 [†] | 9.85 [§] |
| (Split) Before | | 3,088,170 | 50.3 | 6.82 | 6.82 | 6.82 |
| (Split) After | | 827,823 | 13.5 | 19.54 | 18.99* | 11.87 [§] |
| Total | | 6,137,671 | 100.0 | 10.56 | 10.34 | 8.90 |

Sources and notes: TSE (2020) and author's calculations.

*: data adjusted by precinct.

†: data adjusted by locality.

‡: data otherwise adjusted.

§: "Villegas Estimate," based on early results in Villegas's three geographies, weighted by actual votes cast.

Villegas clearly underestimates by a large margin Morales's support. In *split* precincts that account for 84 percent of the post-interruption vote, Villegas underestimates by more than 7.6 percentage points.

We may break down the results further by each of Villegas's geographies, as seen in **Tables 22–24**. First, we see the results for the capital cities and El Alto. Note that regardless of the type of precinct or locality, Villegas adjusts all margins *after* the interruption to -9.85 percentage points — the weighted average of -22.10 and -3.56 as observed *before* the interruption. In the *split* precincts in these 10 municipalities, we would expect Morales to

perform a full 10 percentage points better on the late vote than Villegas estimates. The difference, of course, is the different weighting of precincts for which Villegas fails to account.

Table 22

“Ciudad” Results: Including Full Adjustment and Villegas’s Estimate

| Precinct Type | Locality Type | Valid Votes | Percent of Total | Absolute Margin | | |
|----------------|---------------|-------------|------------------|-----------------|--|-----------|
| | | | | Official | Precinct*, Locality†, or Otherwise‡-Adjusted | Villegas§ |
| All Before | | 2,067,788 | 33.7 | 9.47 | 9.47 | 9.47 |
| All After | All after | 131,480 | 2.1 | 59.94 | 54.77† | 30.17§ |
| | (split) After | 22,410 | 0.4 | 6.23 | -5.67† | 9.85§ |
| (split) Before | | 3,088,170 | 50.3 | 6.82 | 6.82 | 6.82 |
| (split) After | | 827,823 | 13.5 | 19.54 | 18.99* | 11.87§ |
| Total | | 6,137,671 | 100.0 | 10.56 | 10.34 | 8.90 |

Sources and notes: TSE (2020) and author’s calculations.

*: data adjusted by precinct.

†: data adjusted by locality.

‡: data otherwise adjusted.

§: “Villegas Estimate,” based on early results in Villegas’s three geographies, weighted by actual votes cast.

In most of the rest of Bolivia, the differences are not so stark. A large number of votes in this rural grouping came from localities counted *all after* the interruption, making estimates less reliable. Nevertheless, Villegas underestimates Morales’s support relative to more detailed accounting of results.

Table 23**“Provincia” Results: Including Full Adjustment and Villegas’s Estimate**

| Precinct Type | Locality Type | Valid Votes | Percent of Total | Absolute Margin | | |
|----------------|---------------|-------------|------------------|-----------------|--|--------------------|
| | | | | Official | Precinct*, Locality†, or Otherwise‡-Adjusted | Villegas§ Estimate |
| All Before | | 961,927 | 36.8 | 40.17 | 40.17 | 40.17 |
| All After | All After | 124,940 | 4.8 | 61.42 | 55.77† | 31.12§ |
| | (Split) After | 9,075 | 0.3 | 10.58 | 8.26† | 31.12§ |
| (Split) Before | | 1,103,722 | 42.3 | 23.24 | 23.24 | 23.24 |
| (Split) After | | 412,387 | 15.8 | 35.92 | 34.86* | 31.12§ |
| Total | | 2,612,051 | 100.0 | 33.26 | 32.82 | 31.12 |

Sources and notes: TSE (2020) and author’s calculations.

*: data adjusted by precinct.

†: data adjusted by locality.

‡: data otherwise adjusted.

§: “Villegas Estimate,” based on early results in Villegas’s three geographies, weighted by actual votes cast.

Third, we have results from votes cast outside Bolivia. Nearly 84 percent of these votes came from split precincts where we would have the greatest confidence in the adjusted results. The fact that Villegas underestimates Morales’s margin there by more than 17 percentage points tells us his estimates are under-informed by the data.

Table 24**“Extranjero” Results: Including Full Adjustment and Villegas’s Estimate**

| Precinct Type | Locality Type | Valid Votes | Percent of Total | Absolute Margin | | |
|----------------|---------------|-------------|------------------|-----------------|--|--------------------|
| | | | | Official | Precinct*, Locality†, or Otherwise‡-Adjusted | Villegas§ Estimate |
| All Before | | 114,823 | 56.9 | 24.66 | 24.66 | 24.66 |
| All After | All After | 3,541 | 1.8 | 27.51 | 27.60† | 30.25§ |
| | (Split) After | 1,735 | 0.9 | 58.97 | 61.34† | 30.25§ |
| (Split) Before | | 54,592 | 27.1 | 42.01 | 42.01 | 42.01 |
| (Split) After | | 26,938 | 13.4 | 47.65 | 47.23* | 30.25§ |
| Total | | 201,629 | 100.0 | 32.77 | 32.74 | 30.25 |

Sources and notes: TSE (2020) and author’s calculations.

*: data adjusted by precinct.

†: data adjusted by locality.

‡: data otherwise adjusted.

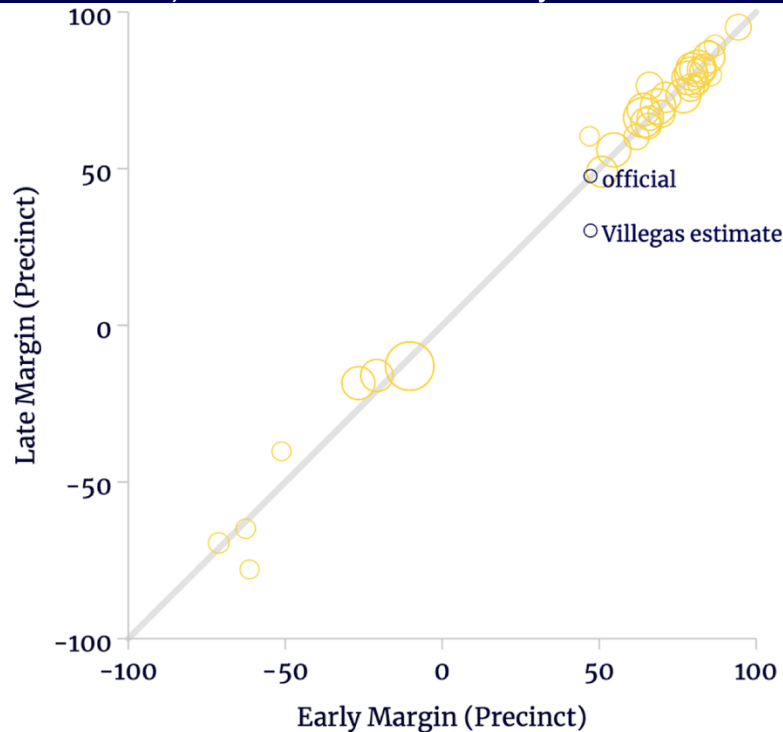
§: “Villegas Estimate,” based on early results in Villegas’s three geographies, weighted by actual votes cast.

We see again the difference in **Figure 4**, which shows the official margins *before* and *after* the interruption for each *split* precinct outside Bolivia. The late margin for each precinct is very

close to what we would expect based on the early margin in that precinct. It is clear that in the aggregate, the early votes from outside Bolivia more heavily favor Morales because the tally sheets came from Morales–favored areas. The tendency for Morales–favored areas (even outside Bolivia) to be counted disproportionately late strikes again, accounting for Villegas’s underestimate of Morales’s support.

FIGURE 4

Outside Bolivia, Late Results Followed Early Results in Each Precinct



Source: TSE (2020) and author’s calculations.

Chumacero, too, finds nearly the same few votes left unexplained as in our fully adjusted results in Table 16. The results presented in **Table 25** below show Chumacero’s full adjustment, including the urban/rural divide and taking into account adjacency within precincts.³⁵

³⁵ See Table 1 for an explanation of adjacencies. Rather than estimating a late tally sheet’s support for each candidate based on the early data at the entire precinct (administrative order 7), Chumacero gives precedence to any early tally sheets neighboring numerically. For example, Chumacero’s estimate of tally sheet 26684 in Colegio Martin Cardenas (Puerto Chaguaya, La Paz) is based on sheets 26683 and 26685, but not 26686 or 26687.

Table 25**Election Results: Official and Chumacero's Full Estimate**

| Election Results: Official and Challenger Split Estimate | | | | | |
|--|---------------------|-------------|------------------|----------|--------------------|
| | | Valid Votes | Percent of Total | Margin | |
| | | | | Official | Estimated |
| Pre-interruption | | 5,155,958 | 84.0 | 7.88 | 7.88 |
| Post-interruption | In quick-count | 733,962 | 11.9 | 25.93 | 24.65 ^A |
| | Official count only | 247,751 | 4.1 | 20.87 | 20.38 ^B |
| Total | | 6,137,671 | 100.0 | 10.56 | 10.39 |

Sources and notes: TSE (2020), Chumacero (2019), and author's calculations. Note A: Official, less 1.28 percentage points, per Chumacero's "Dif 2" in Cuadro 7. Note B: Official, less 0.49 percentage points, per Chumacero's "Dif 2" in Cuadro 5.

However, Chumacero also presents results without controls for within-locality variation. Again, there is no real reason for presenting this result except to observe that within-locality variation is important to explain the post-interruption quick-count results.

Table 26**Election Results: Official and Chumacero's Limited Estimate**

| | | Valid Votes | Percent of Total | Margin | |
|-------------------|---------------------|-------------|------------------|----------|--------------------|
| | | | | Official | Estimated |
| Pre-Interruption | | 5,155,958 | 84.0 | 7.88 | 7.88 |
| Post-Interruption | In Quick Count | 733,962 | 11.9 | 25.93 | 19.03 ^A |
| | Official Count Only | 247,751 | 4.1 | 20.87 | 19.25 ^B |
| Total | | 6,137,671 | 100.0 | 10.56 | 9.67 |

Sources and notes: TSE (2020), Chumacero (2019), and author's calculations. Note A: Official, less 6.90 percentage points, per Chumacero's "Dif 1" in Cuadro 7. Note B: Official, less 1.62 percentage points, per Chumacero's "Dif 1" in Cuadro 5.

Likewise, Newman, in his original paper, suggests that a second round would have been necessary, writing:

[t]he Oaxaca-type decompositions allows for a counterfactual estimate of the extent of the influence of the change in the distribution of margins before and after the cut-off for *municipios* that favored CC before the cut-off. The counterfactual estimates suggest that a second round would have had to be held were it not for the change in distributions of margins before and after the cut-off for *municipios* supporting CC before the cut-off.

But Newman's finding depends on attributing the different margins to changes in trend and not at all on changes in the number of valid votes at each precinct within municipalities. Just as ignoring this important source of within-locality variation drives his subsequent results, Newman's municipality-level analysis fails.

Analyses of the Official Count

While much of Chumacero's paper analyzes the preliminary vote count, one of his primary conclusions concerns the official (*cómputo*) count:

Donde sí se encuentran efectos de primer orden es en el comportamiento demasiado atípico de las últimas actas procesadas por el organismo electoral. Se sabe que parte importante de ellas venían de lugares rurales de Potosí y Chuquisaca, que se esperaba fueran más proclives hacia el MAS. Sin embargo, el margen de votación en el último grupo actas procesadas es atípicamente favorable al MAS si se la compara con actas de recintos, localidades, o municipios similares. Aunque el margen a favor del MAS es de más del 40%, nuestras estimaciones sugieren que el margen esperable debió ser sustancialmente menor.

(Where first-order effects can be found is in the exceedingly atypical behavior of the last tally sheets processed by the electoral body. It is known that an important part of them came from rural places in Potosí and Chuquisaca, which would be expected to be more inclined towards the MAS. However, the voting margin in the last group of tally sheets that were processed is atypically favorable to the MAS when compared with tally sheets from similar precincts, localities, or municipalities. Although the margin in favor of the MAS is over 40%, our estimates suggest that the expected margin should have been substantially lower.)

The official count is interesting and illustrates our arguments about the importance of precinct matching.

In Chumacero's estimation, Morales overperformed on the last 20 percent of tally sheets of the official count by 4.6–5.0 percentage points, once accounting for geography. Adjusting for this unexpected result knocks 0.9 percentage points off Morales's margin, enough to change the outcome of the election. Recall that in the preliminary count, Morales's margin is severely underestimated if within-locality variation is not taken into account. The importance of precinct-level analysis is clear. However, in analyzing the official count, Chumacero reports little difference between an analysis at the locality level and another that accounts for differences within localities.

The crux of the problem is that while the preliminary count had a tendency to count tally sheets from Morales-favored precincts later than those which recorded votes from opposition-heavy precincts, the official count was ordered even more strongly in terms of geography. While results from the precincts were transmitted electronically to central computers for processing in the preliminary count, the official count required hand delivery of physical election materials to a few department offices. Thus, the offices first received tally sheets from more urban, opposition-friendly precincts located nearby and last received tally sheets from more remote, difficult-to-reach areas of the country that strongly supported Morales.

The requirement for physical delivery meant that tally sheets from the same precinct were officially counted much more closely in time than they had been in the preliminary count. Hence, precincts were much more likely to be counted either entirely in the first 80 percent or entirely in the last 20 percent. There is nothing special in this regard to the 80/20 break. Tally sheets from a given precinct are also likely to be all counted before the 38th percent of the count's progression or all counted after the 38th. The critical observation here is that there will be few precincts split across any given point in the count.

We have already seen that when we look beyond individual precincts, prediction can get unreliable because voting within precincts is much more uniform than voting between precincts. This means that the more the late vote comes from *split* precincts, the more reliably we can project those votes.

In **Table 27**, we see that in the official count, tally sheets in the last 20 percent of the official count that come from *split* precincts are very few (4.2 percentage points out of 19.3 percent of all votes). By contrast, in the preliminary count tally sheets counted late from precincts *split* by the interruption in reporting were more the rule than the exception (13.5 percentage points out of 16 percent of all votes). The more we rely on results from outside a given precinct, the more the analysis depends on blind faith that those results are representative.

Table 27

| Overall Election Data Categories: Comparison of Counts | | | | | |
|--|---------------|-------------------|------------------|----------------|------------------|
| Precinct Type | Locality Type | Preliminary Count | | Official Count | |
| | | Valid Votes | Percent of Total | Valid Votes | Percent of Total |
| All Before | | 2,067,788 | 33.7 | 3,561,282 | 58.0 |
| All After | All After | 131,480 | 2.1 | 593,076 | 9.7 |
| | (Split) After | 22,410 | 0.4 | 331,387 | 5.4 |
| (Split) Before | | 3,088,170 | 50.3 | 1,394,984 | 22.7 |
| (Split) After | | 827,823 | 13.5 | 256,942 | 4.2 |
| Total | | 6,137,671 | 100.0 | 6,137,671 | 100.0 |

Sources: TSE (2020) and author's calculations.

Blindly accepting the assumption that the early results in early geographies are representative of the uncounted tally sheets, we show three sets of projections in **Table 28**. The first is a Newman-type adjustment at the precinct level. Because only the few tally sheets counted after the 80 percent mark from *split* precincts are adjusted, the projected results are nearly identical to the official results of the count.

Table 28

| Overall Election Results: Projections Based on the First 80 Percent of Tally Sheets | | | | | | | |
|---|---------------|-------------|------------------|----------|--------------------|--------------------|---------------------|
| Precinct Type | Locality Type | Valid Votes | Percent of Total | Margin | | | |
| | | | | Official | Newman (Precinct) | Chumacero (Full) | Chumacero (Limited) |
| All Before | | 3,561,282 | 58.0 | -0.26 | -0.26 | -0.26 | -0.26 |
| All After | All After | 593,076 | 9.7 | 52.52 | 52.52 | 46.78 | 46.78 |
| | (Split) After | 331,387 | 5.4 | 34.10 | 34.10 | 24.38 [†] | 24.38 [†] |
| (Split) Before | | 1,394,984 | 22.7 | 12.12 | 12.12 | 12.12 | 12.12 |
| (Split) After | | 256,942 | 4.2 | 24.93 | 24.54 [*] | 24.54 [*] | 24.54 [*] |
| Total | | 6,137,671 | 100.0 | 10.56 | 10.55 | 9.44 | 9.40 |

Sources and notes: TSE (2020) and author's calculations. Numbers differ from Chumacero because his method of ordering tally sheets in the official count is unclear.

*: data adjusted by precinct.

†: data adjusted by locality.

The second projection is a full geographic adjustment where all late tally sheets are adjusted based on the smallest available geographies (*split* precincts based on early precinct data, *split* municipalities based on early municipality data, etc.). Just as with the preliminary count, the adjustments in precincts counted all after the 80 percent mark severely underestimated the official result.

Whereas we could not concern ourselves too much with the underestimate of Morales's support in precincts where votes were all counted after the interruption in the reporting of preliminary results given that there were relatively few votes at stake, this simply is not the case in the official count. Imagine an election in the United States in which all the votes from Democratic-heavy New York, Massachusetts, and California were counted first. We would be ill-advised to naively project the vote in the Republican strongholds of Alabama and Wyoming based on these early results. It doesn't much matter if we try to project by precinct or county when the results we have are almost all from other states.

The third projection is like the second, but within-locality variation is ignored. Because this limitation affects only 4.2 percent of votes, it has little impact on the results. Nevertheless, the estimate is clearly inferior, underestimating Morales's support on late tally sheets from *split* precincts by another 0.84 percentage points.

What can we do to improve our estimate of the last 20 percent of the official count if we lack precinct-level data? We could try to make inferences based on 2016 referendum data, as we informally demonstrated above. However, we have a very accurate source of information as well: the preliminary count.

By the time the official count reached 80 percent, the preliminary count was nearly complete. We may exploit the fact that the two counts were ordered differently by accepting as officially reported any tally sheets included in the first 80 percent of the official count *or* reported prior to the interruption of the reporting preliminary count. In **Table 29** below, we repeat the above analysis, but "after" refers to tally sheets in the last 20 percent of the official count *and* after the interruption in reporting. This gives us a much larger collection of tally sheets upon which we may draw. The predictions, correspondingly, are quite close to the official results.

Table 29**Overall Election Results: Projections Based on the First 80 Percent of Tally Sheets and Pre-Interruption Tally Sheets**

| Precinct Type | Locality Type | Valid Votes | Percent of Total | Margin | | | |
|----------------|---------------|-------------|------------------|----------|--------------------|--------------------|---------------------|
| | | | | Official | Newman (Precinct) | Chumacero (Full) | Chumacero (Limited) |
| All Before | | 4,792,629 | 78.1 | 4.78 | 4.78 | 4.78 | 4.78 |
| All After | All After | 65,659 | 1.1 | 65.99 | 65.99 | 58.98 | 58.98 |
| | (Split) After | 1,532 | – | 7.38 | 7.38 | –6.85 [†] | –6.85 [†] |
| (Split) Before | | 1,089,386 | 17.7 | 27.78 | 27.78 | 27.78 | 27.78 |
| (Split) After | | 256,942 | 3.1 | 38.77 | 38.14 [*] | 38.14 [*] | 38.14 [*] |
| Total | | 6,137,671 | 100.0 | 10.56 | 10.55 | 10.47 | 10.35 |

Sources: TSE (2020) and author's calculations.

*: data adjusted by precinct.

†: data adjusted by locality.

Again, adjusted results from split localities or larger geographies underestimate the official results. And again, ignoring within-locality variation causes a further underestimate of the official results. In both cases, however, the number of votes impacted is low. Simply, by the time the official count included 80 percent of tally sheets, we had an enormous amount of information. At that point, nearly 96 percent of tally sheets had been counted one way or another. More than three-quarters of the vote had been in precincts for which all tally sheets had been reported.

In case there is any lingering doubt about those poor estimates in large geographies, we can simply project the first 20 percent of the official count instead of the last. Because the projections will be based on relatively Morales-friendly tally sheets, these overestimate considerably Morales's support whenever variation between precincts is ignored.

Table 30

Overall Election Results: Projections Based on the Last 80 Percent of Tally Sheets

| Precinct Type | Locality Type | Valid Votes | Percent of Total | Margin | | | |
|----------------|---------------|-------------|------------------|----------|---------------------|---------------------|---------------------|
| | | | | Official | Newman (Precinct) | Chumacero (Full) | Chumacero (Limited) |
| All Before | | 4,093,726 | 66.7 | 19.23 | 19.23 | 19.23 | 19.23 |
| All After | All After | 167,201 | 2.7 | 10.68 | 10.68 | 18.51 | 18.51 |
| | (Split) After | 329,009 | 5.4 | -28.04 | -28.04 | -19.47 [†] | -19.47 [†] |
| (Split) Before | | 786,110 | 12.8 | 2.94 | 2.94 | 2.94 | 2.94 |
| (Split) After | | 761,625 | 12.4 | -11.49 | -11.71 [*] | -11.71 [*] | -5.51 [*] |
| Total | | 6,137,671 | 100.0 | 10.56 | 10.54 | 11.21 | 11.98 |

Sources and notes: TSE (2020) and author's calculations. Numbers differ from Chumacero because his method of ordering tally sheets in the official count is unclear.

*: data adjusted by precinct.

†: data adjusted by locality.

We would not suggest this is in any way indicative of fraud. Morales's poor performance on the first fifth of the official count simply reflects that we are making unrealistic assumptions about a sample (the last 80 percent) of tally sheets. As always, projections from *split* precincts do not differ by much.

Finally, we turn again to Nooruddin, who argues that the results in the last 5 percent of the official count were unexpected. Nooruddin's analysis for the OAS was — until recently, when he at long last shared his data — irreproducible. As it happens, Nooruddin's results depend on two factors. First, he orders tally sheets by their time stamp of the second transcription without regard to whether the tally sheet is set aside for review. This results in a different sequence than what appears in the public data. Second, and more importantly, Nooruddin does not sort those time stamps chronologically, but alphanumerically. That is, a time stamp of 2:04 p.m. on a given day came “before” 2:05 a.m. and 12:04 p.m. Nooruddin's analysis simply does not reflect any real-world progression of the count whatsoever. Because he sought to analyze trends over time, this invalidates his entire examination of the official count.

There are actually other data problems with his analysis. One, the data set presents missing values in place of zeros. This means, for instance, Nooruddin removes from his trends and tables of MAS vote shares any tally sheet where Morales received zero votes. Because these tend to appear early in the count, this exaggerates the increase in MAS share over time.³⁶

³⁶ Because his analysis of the preliminary count uses the results of the official count (as do we and as do other researchers) this problem carries over to his analysis of the preliminary count as well. The time stamp issue is

Second, while Nooruddin does note that “Potosí” is recorded as “PotosÃ” for tally sheets not verified in the preliminary count, he fails to take this into account in his official count tables.

Data errors of this sort are not unusual, but it is to the discredit of the OAS that despite months of skeptical inquiry from researchers, no effort has been made to fix these problems.

Again, we emphasize that if the last 5 percent of either count (suddenly or slowly) were to favor Morales, this would not be evidence of fraud. It would not even rise to the level of suspicion unless it were genuinely inexplicable. However, Nooruddin greatly exaggerates the extent of the change. Below, we produce two tables. In the first, we reproduce the OAS table at the top of page 93 of the audit report. We do not correct for Nooruddin’s data errors, but we compare the last 5 percent (according to Nooruddin) to the prior 5 percent. The second is identical to the first, but with Nooruddin’s time stamps sorted chronologically.

Table 31

MAS and CC Vote Shares in Last 10 Percent of Official Count (Alphanumerical Sort)

| | PS-Level MAS Vote Share | | PS-Level CC Vote Share | | MAS Advantage Over CC | |
|------------|-------------------------|--------|------------------------|--------|-----------------------|--------|
| | 90–95 | 95–100 | 90–95 | 95–100 | 90–95 | 95–100 |
| National | 51.5 | 59.7 | 25.8 | 18.9 | 25.7 | 40.9 |
| Beni | 34.3 | 39.2 | 33.0 | 23.5 | 1.3 | 15.7 |
| Chuquisaca | | 63.7 | | 19.0 | | 44.7 |
| Cochabamba | 55.3 | 71.1 | 31.9 | 16.4 | 23.4 | 54.7 |
| La Paz | 58.2 | 53.6 | 15.2 | 18.9 | 43.0 | 34.7 |
| Potosí | | 71.6 | | 8.7 | | 63.0 |
| Santa Cruz | 41.6 | 50.6 | 33.4 | 25.4 | 8.2 | 25.2 |

Sources: TSE (2020), Nooruddin (2020), and author’s calculations.

not a problem in the preliminary count because those time stamps happen to be formatted in a way that sorts identically both alphanumerically and chronologically.

Table 32

| MAS and CC Vote Shares in Last 10 Percent of Official Count (Chronological Sort) | | | | | | |
|--|-------------------------|--------|------------------------|--------|-----------------------|--------|
| | PS-Level MAS Vote Share | | PS-Level CC Vote Share | | MAS Advantage Over CC | |
| | 90-95 | 95-100 | 90-95 | 95-100 | 90-95 | 95-100 |
| National | 60.2 | 62.8 | 16.3 | 17.0 | 43.9 | 45.8 |
| Beni | | 51.3 | | 19.7 | | 31.6 |
| Chuquisaca | | 63.7 | | 19.0 | | 44.7 |
| Cochabamba | 63.3 | | 23.4 | | 39.9 | |
| La Paz | 60.2 | 59.1 | 15.9 | 18.9 | 44.3 | 40.2 |
| Potosí | | 71.6 | | 8.7 | | 63.0 |
| Santa Cruz | | 37.2 | | 35.6 | | 1.6 |

Sources: TSE (2020), Nooruddin (2020), and author's calculations.

Rather than a 15 percentage point increase in Morales's "advantage" over Mesa, a chronological analysis finds an increase of less than 2 percentage points. Nooruddin's analysis of the final 5 percent of the official count is simply wrong, and this does not even take into consideration the effects of data errors or whether the increase might be predictable.

In **Table 33**, we project Morales's margin at the 95 percent mark. At the point in the official count when 95 percent of valid votes had been verified, Morales held an 8.61 percentage point lead over Mesa. Of the 5 percent of valid votes outstanding, 3.8 percentage points (75 percent) were on tally sheets that had been presented in the preliminary count, prior to the interruption of reporting. Morales won these by 46 percentage points. This brought his margin up past the 10 percentage point threshold for victory in the first round.

Table 33

| Overall Election: Adjusting the Last Uncounted, Unverified Tally Sheets in Official Count | | | | | | |
|---|--------------------------|---------------|-------------|------------------|----------|----------|
| Official Count Status | Preliminary Count Timing | Precinct Type | Valid Votes | Percent of Total | Margin | |
| | | | | | Official | Adjusted |
| Verified | | Before | 5,830,683 | 95.0 | 8.61 | 8.61 |
| | Pre-Interruption | Before | 230,182 | 3.8 | 46.01 | 46.01 |
| Unverified | Post-Interruption | All After | 36,034 | 0.6 | 67.94 | 63.77 |
| | | (Split) After | 40,772 | 0.7 | 39.93 | 38.24 |
| Total | | | 6,137,671 | 100.0 | 10.56 | 10.53 |
| <i>Unadjusted tally sheets (verified or pre-interruption)</i> | | | | | 10.03 | 10.03 |

Sources: TSE (2020), Nooruddin (2020), and author's calculations.

Though Morales would not require it for victory, the few tally sheets remaining were clearly going to favor Morales by a significant amount. A final Morales margin of more than 10.5 percentage points could be expected from the 95 percent of tally sheets that had been verified in the official count. In short, the late increase in Morales's margin was predictable, and in any case small.

Thus, while Morales's margin increased considerably late in the official count (whether the last 5 percent, the last fifth, or even the first fifth) it is clear that the importance of precinct-level analysis is critical for making accurate predictions about the remaining votes. It is absurd to argue that the final result is inexplicable without taking this into account.

Further Rebuttal to Newman

Much of Newman's response to our critique danced around the question of precinct-level analysis. We have made abundantly clear that analysis reliant upon wider administrative divisions such as localities, municipalities, and departments systematically understates Morales's expected margin because within these larger geographies Morales's support was counted disproportionately late. It is important to reduce the effect of this bias by making use of early precinct data or by reverting to information gleaned from prior elections. Newman's original municipality-level analysis is lacking in this regard; his subsequent locality-level analysis is hardly an improvement. However, there are other problems with Newman's response.

Unlike our analyses above, Newman groups tally sheets into geographies that either favored or disfavored Morales relative to Mesa prior to the interruption. Rather than adjusting all the tally sheets after the interruption, Newman adjusts only those in geographies that disfavored Morales early. In other words, he does not project the tally sheets where Morales can be expected to win — only those where he can be expected to lose.

We argue that there is no good reason for this differing treatment of precincts based on early results. Newman, however, tries to justify not adjusting Morales-favored areas based on the suspicion that the early and late margins are not demonstrably different.

Suppose, for example, that in a particular precinct Mesa wins overall 105 to 95, and we take 20 votes as a random "late" sample of that precinct. As the sample is random, the "early" and

“late” votes are identical in every way, within sampling error. Just as a survey of voters has a margin of error, so our late sample may not be precisely representative.

In fact, we will find Morales will win on that “late” sample more than 30 percent of the time, though he lost on the 180 “early” votes. In any precinct where Mesa won early, there is (obviously) a 100 percent chance that Mesa won early, but there is some chance that he also lost late. In technical language, conditioning limits the support of the distribution of early margins (the early “Mesa” margins must be negative) but not that of the late margins (the late “Mesa” margins are mostly negative but sometimes positive.) With different support, the distributions are inherently different.

Sampling error alone implies that the early and late margins *conditional on Mesa winning early* have different distributions.

A statistical test is therefore unnecessary. If sufficiently sensitive, a test should detect this difference — a difference that will exist even in the complete absence of any fraud whatsoever — and report a true positive. In practice, Newman’s test is not sufficiently sensitive and will frequently report a false negative.

Thus, Newman’s test cannot be used to identify suspicious votes; once Newman divides the sample into “won early” and “lost early,” the test can only report *accurately* a positive result. However, Newman misstates our criticism of his approach to testing the election data, writing:

If this result is universally true, *if, regardless of the data involved, the hypothesis of equal distributions is always rejected when one splits the sample*, then it should have been impossible for Newman (2020) to fail to reject the hypothesis of equal distributions for the simulated cut-off applied to the set of recintos that had all their votes reported before the actual cut-off (as described in Annex B of Newman, 2020). *[Emphasis added.]*

Newman’s premise is not ours. We do not argue that “the hypothesis of equal distributions is always rejected,” but that “the hypothesis of equal distributions is always false.” There are four possibilities, as in **Table 34**.

Table 34

Hypothesis: Early and Late Distributions Are “Equal”

| | | Test Result | |
|--|-------------------------|--|--|
| | | Cannot say distributions are not “equal” (“fail to reject hypothesis”) | Distributions are probably not “equal” (“reject hypothesis”) |
| Are the distributions actually “equal” in reality? | Yes (“true hypothesis”) | True Negative | False Positive |
| | No (“false hypothesis”) | False Negative | True Positive |

The two rows correspond to different worlds where the hypothesis is alternately true (first row), or false (second row). The two columns correspond to different possible test results within each world: negative (first column), and positive (second column). The purpose of testing is to help us figure out which world we are in. A perfect test would only yield outcomes in blue. If the distributions are equal, the test result will be a true negative and we would know the distributions are equal; if unequal, the result will be a true positive and we would know the distributions are unequal. Unfortunately, no test is perfect, and we must account for the orange outcomes as well as the blue.

In any case, Newman argues that if we lived in a world restricted to the second column, then it should be impossible to find a negative result. This is obviously true, but our argument is that we are in a world corresponding to the second *row*. Negative results are possible in this world. Rewriting Newman to reflect our argument:

If this result is universally true, if, regardless of the data involved, the hypothesis of equal distributions is always *correctly* rejected when one splits the sample, then it is impossible for Newman (2020) to *accurately* fail to reject the hypothesis of equal distributions for the simulated cut-off applied to the set of recintos that had all their votes reported before the actual cut-off (as described in Annex B of Newman, 2020). *[Edits in emphasis.]*³⁷

We agree that Newman should be able to “fail to reject the hypothesis,” i.e., get a negative test result. Such a result is simply a false negative.

In fact, we argued that when Newman divides geographic areas into “won early” and “lost early,” the tests in effect measure the percentage of geographic areas where the candidate

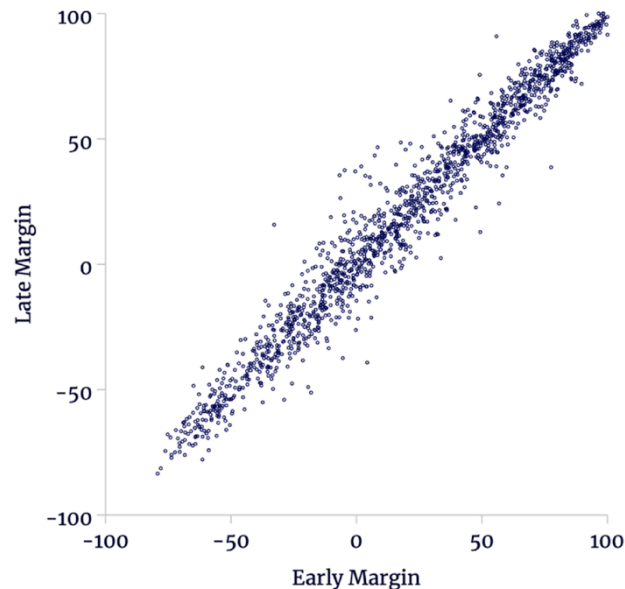
³⁷ Newman (2020a).

that won late differed from the candidate that won early. Sampling error all but guarantees us that this will happen. Indeed, it would be very suspicious if this did *not* happen.

Let us take a quick look in **Figure 5** at the precincts that were partially but not completely counted prior to the interruption in reporting (*split* according to Newman’s definition).

FIGURE 5

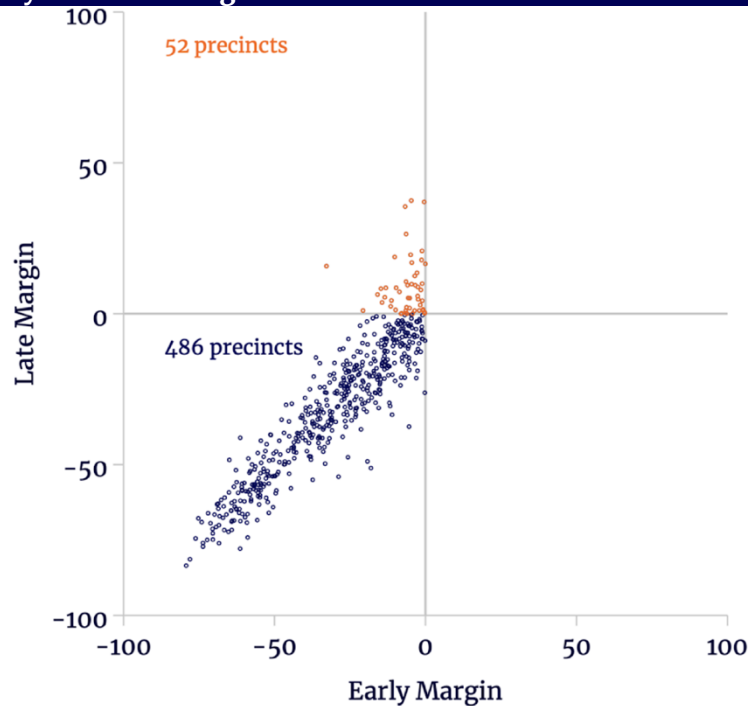
Early and Late Margins in All Precincts



Source: TSE (2020), Newman (2020b), and author’s calculations.

Clearly, there are a few outliers, but they seem not to favor either candidate. To some extent, several of these differences between early and late margins are due to low sample size. However, this is relatively unimportant.

In **Figure 6**, we look only at the “Mesa” precincts — those precincts where, in the early part of the count, Mesa received more votes than Morales.

FIGURE 6**Early and Late Margins in “Mesa” Precincts**

Source: TSE (2020), Newman (2020b), and author's calculations.

We see that in 486 of the 538 precincts considered here, Mesa received more votes than Morales on the tally sheets counted late. However, this was not true for 52 precincts (upper left quadrant). It is critical to note that 52 of 538 is 9.67 percent — exactly the test statistic reported by Newman in his original Table 6. There is not an exact correspondence between the two, but it is not a coincidence.

By construction, 100 percent of the early margins among the 538 are below 0. However, only 90.33 percent of the late margins are below 0. Some of the 52 precincts generally favored Mesa by a small margin overall, but happened by chance to favor Morales on a few late tally sheets. Others favored Morales by a small margin overall, but happened by chance to favor Mesa in a few early tally sheets. Again, we say the margins in these 52 precincts “bleed” past the 0 cutoff. That this happens is natural, and is not a sign of fraud. Most importantly for our purposes, Newman’s Kolmogorov–Smirnov (K–S) test statistic D (9.67) is picking up quite literally the share of precincts that bleed.

Somewhat informally, if very technically, D is a measure of the maximum difference in cumulative distribution functions. If, for example, 100 percent of “early” observations are negative, but only 90 percent of the “late” observations are negative, then D must be at least 10 — the difference between 100 and 90. Depending on the number of observations, a D of

10 may be very large, or relatively small. If D is determined to be large, the K-S test is positive; it rejects the hypothesis that the distributions are equal.

The connection to bleeding precincts is immediately clear. In our example, 100 percent of the early margins are negative, but only 486 of 538 *split* precincts have negative late margins. The difference — the percentage that bleed — is 9.67 percentage points, so D is at least 9.67 percent.

In this case, the statistic is exactly that number. In general, if the probability of a precinct bleeding due to sampling error is d , then D must be at least d , regardless of the number of precincts observed. However, the critical value of D — the threshold for deciding between a positive or negative result — falls with the number of precincts. The more precincts we observe with a relatively steady share d of bleeding precincts, the more likely we are to find a positive test result.

This implies that a negative result here is a consequence of insufficient observations — or a test that lacks sufficient sensitivity. It is not, as Newman argues, because there is something suspicious about the data itself. We will illustrate this with simulated data, but we must first address Newman’s concern regarding our choice of simulation.

Newman makes hay of the fact that our simulated data does not match the actual data. However, this is irrelevant. As we concluded in our original response: “Newman’s approach to discerning fraud fails when applied to clean data, so there is no reason to trust his test results when applied to the actual election results.”³⁸

The point of our simulations was to illustrate the general problem with splitting the data in the way Newman has. Newman seems to argue that though his approach fails on one simulated type of “clean” election, his approach is still valid for another, but he does not explain why his approach should not generally apply. It is a curious sort of test that, as Newman argues, makes sense for the results of this one election and this one election only.

Newman does propose an alternative approach to generating simulated data “to match the actual data of interest.” Yet he acknowledges that this fails:

³⁸ Rosnick (2020).

The alternative simulated data set does poorly in matching the votes for MAS and CC after the cut-off. It underpredicts the MAS votes by 12.2 percent and overpredicts the CC votes by 17.6 percent.

These differences are very large. In contrast to the actual data, where the Morales-favored tally sheets tended to be counted later than those that favored Mesa, Newman's data permits no such tendency. He writes: "There are no compositional changes in Recintos at play." This applies to our own simulations as well, but it is the lack of compositional change across (not within) precincts between the early and late counts that causes his calibration to fail.

For now, however, we are happy to accept Newman's process for generating simulated precincts as valid in order to thoroughly examine Newman's test.

To begin, note that Newman argues that on a sample of 1,496 observations generated by his approach, his tests return all negative results (suggesting that the early and late distributions are not distinct). But again, this finding does not address our argument. All negative results on split data are false negatives. The fact that Newman is able to generate negative results therefore proves nothing. Simulated or otherwise, it would be very strange if, among precincts where one candidate won early, the other candidate could never win late but for fraud. Because of sampling error, this simply makes no sense. The question is whether or not the rate of bleed is unexpectedly high.

Fortunately, Newman's method of generating simulated data makes the bleed rate easy to calculate. We find that 7.3 percent of precincts with an early margin below 0 (2.9 percentage points of 39.3 percent of all precincts) will bleed and have a positive margin late. If we increased the sample size to 1,745, we would expect 685 "Mesa" precincts. The critical value of D for 685 observations is 0.073, meaning the K-S test would, in effect, be no better than a coin flip.

For only 545 "Mesa" precincts — a smaller sample size — the critical value of D is even larger, meaning a false negative is more likely than a true positive.

With an eye to quality control, we may apply Newman's chosen test to a large number of simulated data sets, all drawn according to his specification. If the test is appropriate and Newman's interpretation of his simulated results is correct, then we should find "false" positives in 5 percent of the draws, regardless of the number of observations generated. This is what it means for a test of significance at the 5 percent level: the critical value is (or table

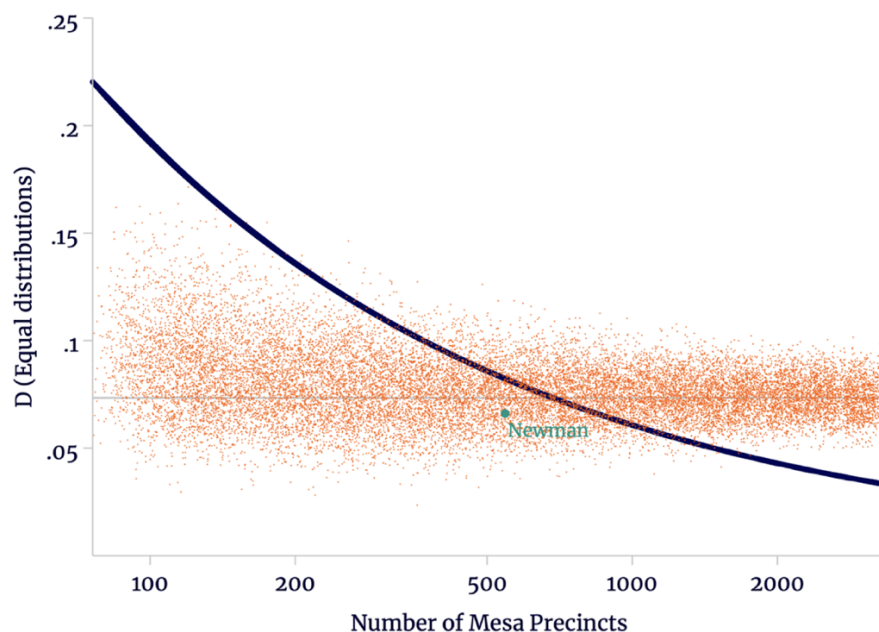
of critical values are) chosen to permit a false positive rate of 5 percent. In fact, we get a “false” positive 337 times in 1,000 different random draws of 1,496 simulated precincts each. This has two implications: first, the critical value was too low for the chosen level of significance; second, the single random draw employed by Newman was not entirely representative of his process for data generation. In other words, though odds were in his favor, he was fortunate nevertheless to find a result consistent with the story he tells in his response.

However, as we noted above, the likelihood of finding a positive result ought to vary with the number of observations. As this is simulated data, we have no problem observing this.

Figure 7 shows Newman’s K-S tests of the “Mesa” precincts for 21,000 simulations of varying size. The computed test statistics are shown in orange. Regardless of the number of observations in each simulation, the *D* statistic varies around 7.3 percent — the expected percentage of “Mesa” precincts that bleed. The dark blue line shows the critical values of the test statistic. Test statistics below the line indicate a negative result; above the line, positive. Figure 3 also indicates Newman’s reported test result.

FIGURE 7

K-S Test Results for Simulated Data



Source: Newman (2020b) and author’s calculations.

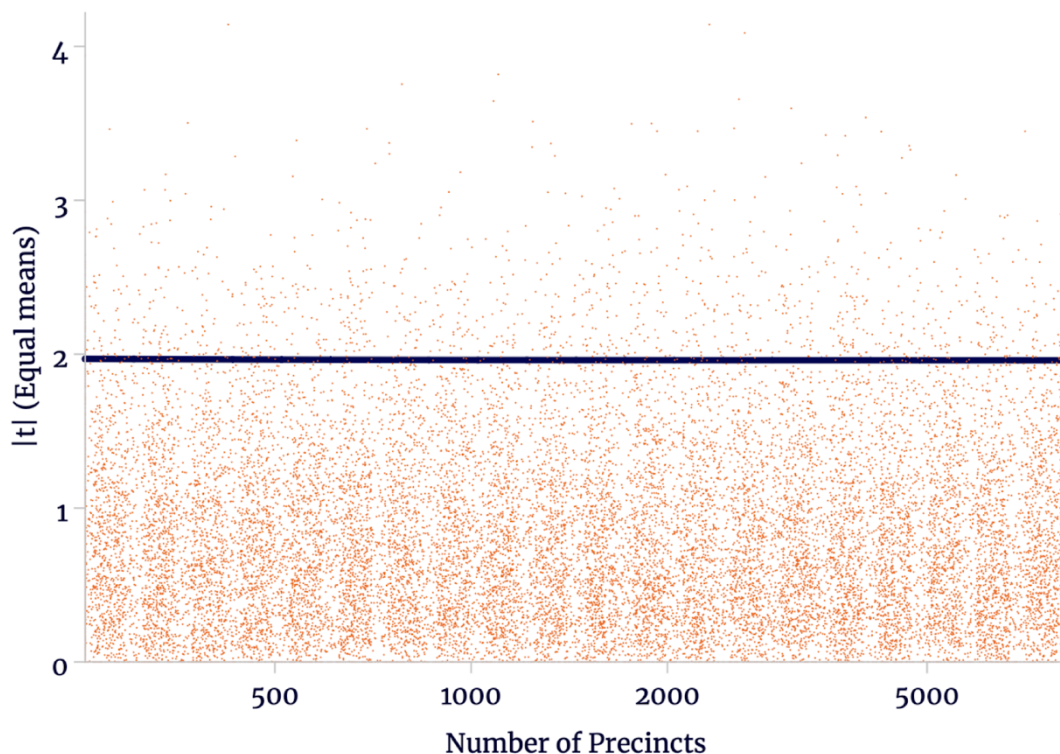
Clearly, while Newman’s reported negative result was not unlikely, it occurs only about twice in every three simulations. For smaller data sets, the test result is consistently negative; for

larger, consistently positive. The fact that the test results (with sampling error) are consistently around the theoretical bleed rate confirms again that it is the percentage of precincts that bleed driving the test results. The fact that the test statistic does not fall as the sample size grows indicates that the simulated early and late margins do in fact differ; only the lack of test sensitivity at the lower sample sizes accounts for the false negative test results.

For contrast, we also observe the results of a less-controversial test of the simulated data. Newman's data generation produces a set of early and late margins, each distributed normally with identical means and variances. This makes a paired t-test for equality of means absolutely ideal, with well-known properties. Because the early and late distributions are known to have identical means (apart from sampling error), any positive result (suggesting differing means) must be false. We should see a consistent 5 percent false positive rate regardless of the sample size. In **Figure 8**, we see exactly this.

FIGURE 8

t-Test Results for Simulated Data



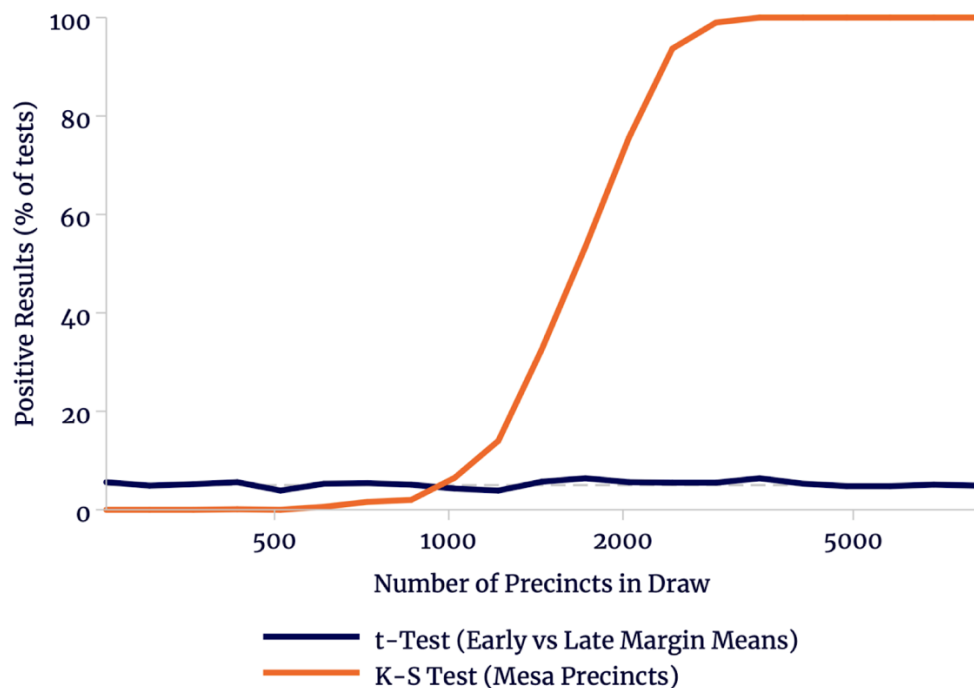
Source: Newman (2020b) and author's calculations.

Finally, we compare Figure 7 and Figure 8 by showing only the likelihood of a positive result for each test. In **Figure 9**, we see how the positive rates increase with sample size. In contrast

to the t-test with true negative results, the K-S test results are inconsistent with the hypothesis that the simulations produce true negative results. Newman simply happened to find a negative result on a single simulation. Had he chosen a different seed for his random numbers, he may well have found a different result; had he chosen a larger sample size, his chances of finding a positive result would have increased considerably. This only makes sense if the negative results of the K-S test are indeed false.

FIGURE 9

Comparing K-S and t-Test Positive Rates



Source: Newman (2020b) and author's calculations.

We therefore conclude that Newman's proposed simulations reinforce our argument that his negative test results on the actual election data are in fact false. The observed distributions are different, even in the absence of fraud. Thus, Newman's tests are uninformative and should be abandoned. More importantly, what Newman identifies as suspicious in the election data is not inherent in the data, but something he creates by conditioning his samples based on the early results.

Having argued that Newman's K-S tests pick up the percentage of geographies that bleed, we proposed a workaround: separating pro-Morales geographies from pro-Mesa geographies by the sum of early and late margins, rather than by the early margin alone. This is by no means a perfect solution, but it has the advantage of simplicity, and maintains the spirit of Newman's division. Geographies with negative margins both early and late will be "Mesa"

precincts, just as in Newman; geographies with non-negative margins both early and late will be “Morales” precincts, just as in Newman. Only geographies that bled under Newman may be categorized differently.

Importantly, early and late margins are treated symmetrically, so the support for both the early and late margins — even when divided — is identical. Early and late margins for both “Mesa” and “Morales” precincts may range from -100 to +100. In practice, sampling error will neither permit “Mesa” margins to stray far into positive territory, nor permit “Morales” margins to stray far into negative territory. However, softening the cutoff to allow for it makes the support for early and late margins identical. This is not sufficient to ensure that fraud-free results will have identical distributions, but it is necessary.

Applying the K-S tests to data divided by this soft cutoff, we found negative results much more often, suggesting that Newman’s results are indeed driven by his cutoff rather than an irregularity in the data.

Newman’s response to our suggestion is entirely nonresponsive. His criticism, such as it is, follows two points. First, that the use of a soft cutoff results in “drastically different” results. This is, of course, our point: Newman’s results are driven by the fact that his division of the data *creates* bleed — late margins that fall outside the support of the early margins — and Newman’s K-S tests pick up on this.

Newman finds it “curious” that six observations one way or another can make the difference, but he understates the number of precincts that switch from one side of the cutoff to the other. The left split adds 20 precincts and subtracts 26. Likewise, the right split picks up 26 and gives up 20. A total of 46 precincts out of an average sample size of 748 (6.1 percent) cross over when moving from a hard to a soft cutoff.

Nor should Newman be surprised that the left split changes more than the right does. The crossing precincts represent a smaller share of the MAS-heavy precincts than they represent as a share of the opposition-heavy precincts. Only 4.7 percent of “Morales” precincts in Newman’s simulations should bleed, for an expected *D* statistic 35 percent smaller. However, the sample sizes increase only enough to reduce the standard critical value by 20 percent. Thus, the false negative rate is higher in “Morales” precincts.

Whereas Newman downplays the degree to which a soft cutoff changes the split in the data, Newman actually adds or removes exactly seven precincts, depending on the choice of

alternative location. The degree to which this could possibly matter is sickeningly small. His Table 5 indicates, for example, that in order to shift seven precincts from the right to the left, the cutoff must be moved by six early votes (on net) in the entire precinct. This corresponds to a switch of three votes from Morales to Mesa. From this, Newman expects to draw an important conclusion.

To say the least, we find it odd that he uses the difference in votes rather than the difference as a share of valid votes cast. With his +6 vote cutoff, Esc. Sub. Central la Porfía (Mineros) is defined as favoring Mesa because Morales won there by only 22 to 18 on the early vote. On the other hand, Morales barely eked out a 933–927 early vote win at Colegio Walter Alpire 2do Patio (El Alto). At the latter precinct, Mesa would come from behind to win this “Morales” precinct overall, 1198–1192. The bottom line here is that Newman’s choice of alternative cutoffs is not informative; his alternative at –3 votes seems even less informative.

On the other hand, in our initial response we presented results having moved the cutoff from 0 to a margin of +20 percentage points and found our results to be robust. Attempting to do likewise with the actual data, we find that Newman’s positive result vanishes. Again, we believe this is a consequence of the low sensitivity of the test and not because there is no difference between the distributions.

Even accepting his choice of robustness test, we question his conclusion. As we have amply demonstrated, the test statistic is deeply connected to the percentage of observations that bleed — precincts that are considered supportive of the opposition in the early count, though Morales wins in the late (and vice versa). Newman’s baseline test statistic suggests that 52 of 538 precincts bleed — that Mesa fails to win these precincts in the late vote despite winning them early; this turns out to be exactly correct. If we shift the cutoff by only a small amount, then we expect to gain or lose a few precincts, but we expect the number of precincts that bleed to remain about the same. If the total number of precincts tested does not change much, then we expect the test statistic to also remain stable. Sure enough, the test statistic does not vary by much. Given that he finds a true positive result on the “Mesa” precincts, it is not surprising that he continues to find a true positive. The false negative result on the “Morales” precincts is likewise stable.

Finally, Newman downplays his own work, writing, “A statistical analysis should not be used to judge whether there was fraud. At best, it can suggest that there are potential problems.” We find it hard to disagree. However, no statistical analysis of Bolivia’s 2019 elections that

has withstood any real scrutiny has suggested potential problems of any real consequence, and this is no exception.

He then cites the “227 [tally sheets] found to be questionable.” Annex 8 of the OAS Final Report finds that in a survey of thousands of Morales-heavy tally sheets, 226 were allegedly irregular. In some voting centers, the OAS notes, one person filled out multiple tally sheets. But it is not clear that this is actually an irregularity; nor is it clear that the OAS allegations are even true. The OAS presentation on handwriting has been widely accepted but is highly selective, and as with many other issues regarding this election, the OAS has been unwilling to answer simple questions let alone subject itself to outside examination.

Were there irregularities in the election, in procedure, and in the recording of the vote? Certainly, so it goes with any election of reasonable size. Everyone who asks, receives, and they who seek, find; as Newman notes in his response, the OAS did seek. Critically, the OAS used their flawed statistical analyses to justify subjecting Morales-heavy tally sheets to special scrutiny, in lieu of a random sample. The OAS handwriting analysis starts explicitly with the 125 tally sheets upon which Morales received at least 99 percent of the vote.

When, inevitably, irregularities — no matter how benign — were discovered, the OAS used this discovery to further suggest that the election result was in doubt. Regarding the 226 tally sheets flagged as “observed to include 2 or more tally sheets from the same polling station filled in by one and the same person,” the OAS writes:

As a whole, all of these [irregular tally sheets] account for 38,001 valid ballots, of which the MAS political party secured 91% of the votes; that is, 34,718 votes, almost the number of votes making it possible for Morales to avoid a second round.

The margin of victory of Morales in the first round amounted to about 40,000 votes. Without them, the difference with Mesa would have been lower than 10% and therefore a second round would have been required.³⁹

The comparison of 34,718 votes for Morales on the flagged tally sheets, to the 34,672 net votes by which Morales’s margin exceeded the 10 percentage point hurdle, is entirely irresponsible.

³⁹ OAS (2019b).

First, the OAS's conclusion is not correct, mathematically speaking. If the tally sheets were simply invalidated, this would reduce Morales's margin over Mesa by no more than 34,718 votes, but would also reduce the number of valid votes in the election by 38,001. The latter has the effect of increasing Morales's margin as a percent of valid votes. (Fewer total valid votes implies the margin is larger as a percentage of valid votes.) On balance, Morales would still hold on to a first-round victory.

Second, even if the tally sheets are irregular, and show votes overwhelmingly in favor of Morales, the remedy is not to throw out these votes. As we noted previously,

According to electoral observation standards, missions should comply with domestic law. There is a process in Bolivia's electoral law for the annulment of tally sheets. Article 177 of the electoral law defines the causes for annulment. In the event that a tally sheet is annulled, the election will be reheld on the second Sunday following the election.⁴⁰

Therefore, the annulment of tally sheets matters only if the repolling at the particular stations would significantly affect the outcome of the election. And it is clear that Morales had considerable support at these irrespective of the alleged irregularities.

To take the first example on the list, the OAS flagged the entire precinct Esc. Sarcoma, which broke 514 to 1 for Morales out of 517 valid votes. The OAS implies that all net 513 votes for Morales are in doubt, but in 2014 Morales won 468 to 4 on 477 votes there; in 2016 voters in favor of lifting term limits won 474 to 16. The municipality of Sacaca, which contains Esc. Sarcoma, voted 5,512 to 40 for Morales.⁴¹ We can state with confidence that this precinct voted on net for Morales by the hundreds.

After repolling the stations, it is certain that Morales's first-round victory would hold.

Even accepting the OAS's handwriting analysis, did the OAS investigate the possibility of benign reasons for such findings? For example, Section III of the General Election Rules call for electoral notaries to assist with logistics and operations.⁴² Perhaps a notary assisted (or supervised another) in filling out tally sheets on behalf of someone semi-literate, or with

⁴⁰ Johnston and Rosnick (2020).

⁴¹ This only counts the "rural" parts. That is, it excludes the location of Sacaca within the municipality. Even within "urban" Sacaca, Morales won 912 to 316.

⁴² OEP (2019).

poor handwriting? Absent strong evidence that signatures and fingerprints are forged, it is hard to envision the numbers on the tally are fraudulent — even if the tally sheets are nevertheless “irregular.”

Coming back to Newman’s concluding argument, we rebut as follows: a flawed statistical analysis should not be used to judge whether there was fraud. Nor can it suggest that there are potential problems. Flawed statistical analyses cannot be seen as suggestive. Nor can the discovery of possible irregularities justify the flawed analysis — even on purely practical grounds.

Suppose I produce a study suggesting that drivers of Ford automobiles are a thousand times more likely to break the speed limit compared to drivers of autos manufactured by others, and this causes police to target Fords. When the police come back having ticketed a bunch of Ford drivers, ought I then pat myself on the back because the “forensic evidence” judges these drivers?

It is no different with the Bolivian election. Newman’s deference to the OAS findings does not absolve him of his defense of the OAS’s claims. Those claims are still irresponsible in the light of all we know, and Newman’s report is still erroneous.

Conclusion

Newman employs insufficient geographic controls in his case supporting OAS allegations of fraud in Bolivia's 2019 election. His failure to account for within-precinct variation in his adjustments — the importance of which he had already observed — utterly undermines his conclusions. Further, Newman's continued defense of his work fails to advance his case.

Newman is not alone in this regard; studies that question the results of these elections universally rely on ignoring variation within broad geographies to underestimate Morales's support in the remaining count. Chumacero also presents locality-level estimates to question the election results. Escobari and Hoover actively discount their precinct-level results in favor of municipality — even uncontrolled — estimates. Nooruddin, on behalf of the OAS, manages nothing better than a half-hearted department-level analysis. Villegas offers even less.

There is no question that the late-counted tally sheets in both the quick and official counts broke disproportionately heavily for incumbent Evo Morales. There is also no question as to whether or not Morales would have secured a first-round victory if not for the fact that the late sheets broke more heavily in his favor: he would not. The question is whether or not the shift is sufficiently understood. If we understand why the late votes bumped Morales past the 10 percentage point threshold for a first-round victory, then the will of the voters was usurped in October 2019 when the OAS — acting as election observer — openly disputed, without evidence, the validity of the late votes.

The complication of intra-geographic variation was immediately apparent. We explained this as early as October 22 — the day after the OAS's irresponsible press release, and only two days after the election itself. The OAS release expressed “deep concern and surprise at the drastic and hard-to-explain change in the trend of the preliminary results after the closing of the polls,” without presenting any evidence.⁴³ This claim, which the OAS was to repeat in subsequent statements, formed one of the main pretexts for Morales's forced removal from office on November 10, 2019.

⁴³ OAS (2019a).

References

- Alvarado, Edgar Villegas, and his investigation team. 2019. "Report on Irregularities Found in the Bolivian Election Held on 20 October 2019." November 6.
<https://www.scribd.com/document/436228825/Report-on-Irregularities-Found-in-the-Bolivian-Election-Held-on-20-October-2019>
- Bolivia en la OEA. 2020. "Estudio del Sr. John Newman, economista y líder estadístico que trabajó en el Banco Mundial confirmando los resultados del informe de auditoría de la OEA sobre las elecciones del 20 de octubre de 2019
<https://drive.google.com/file/d/1bu8mMS...>" Twitter, May 1.
<https://twitter.com/boliviaoea/status/1256294524269314049?lang=en>
- Center for Economic and Policy Research (CEPR). 2019. "OAS Should Retract Its Press Release on Bolivian Election, CEPR Co-Director Says." Washington, DC: Center for Economic and Policy Research, October 22. <https://cepr.net/press-release/oas-should-retract-its-press-release-on-bolivian-election-cepr-co-director-says/>
- Chang, Ha-Joon, et al. 2019. "The OAS Has to Answer for its Role in the Bolivian Coup." The Guardian, December 2.
<https://www.theguardian.com/commentisfree/2019/dec/02/the-oas-has-to-answer-for-its-role-in-the-bolivian-coup>
- Chumacero, Rómulo A. 2019. "El camaleón, el mutante y Houdini: Resultados de las elecciones en Bolivia." November 21. <http://rchumace.econ.uchile.cl/papers/cmh.pdf>
- Escobari, Diego and Gary A. Hoover. 2019. "Evo Morales and Electoral Fraud in Bolivia: A Natural Experiment Estimate." Manuscript, November 27.
https://www.researchgate.net/publication/337485965_Evo_Morales_and_Electoral_Fraud_in_Bolivia_A_Natural_Experiment_Estimate
- Human Rights Watch (HRW). 2020. "Justice as a Weapon: Political Persecution in Bolivia." September 11. <https://www.hrw.org/report/2020/09/11/justice-weapon/political-persecution-bolivia>

Idrobo, Nicolás, Dorothy Kronick, and Francisco Rodríguez. 2020. “Do Shifts in Late-Counted Votes Signal Fraud? Evidence from Bolivia.” June 7.

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3621475

Johnston, Jake, and David Rosnick. 2020. “Observing the Observers: The OAS in the 2019 Bolivian Elections.” Washington, DC: Center for Economic and Policy Research, March 10. <https://cepr.net/report/observing-the-observers-the-oas-in-the-2019-bolivian-elections/>

Kitagawa, Evelyn M. 1955. “Components of a Difference Between Two Rates.” *Journal of the American Statistical Association* 50:272, December.

<https://u.demog.berkeley.edu/~jrw/Biblio/Eprints/%20J-L/kitagawa.1955.pdf>

Kurmanaev, Anatoly and María Silvia Trigo. 2020. “A Bitter Election. Accusations of Fraud. And Now Second Thoughts.” *The New York Times*, June 7.

<https://www.nytimes.com/2020/06/07/world/americas/bolivia-election-evo-morales.html>

Long, Guillaume, David Rosnick, Cavan Kharrazian, and Kevin Cashman. 2019. “What Happened in Bolivia’s 2019 Vote Count? The Role of the OAS Electoral Observation Mission.” Washington, DC: Center for Economic and Policy Research, November 8.

<http://cepr.net/images/stories/reports/bolivia-elections-2019-11.pdf?v=3>

Newman, John. 2020a. “Newman Response to Idrobo et al and Rosnick.”

https://drive.google.com/file/d/12mnE6qKW_cVe-nOVI_G8xKxFAQ5X-v55/view

Newman, John. 2020b. “The OAS Conclusions about the Election Integrity of the Bolivian Election are Correct.” April 29.

<https://drive.google.com/file/d/1bu8mMS5ElyjVCkdmTtrCeMb38n1F-w8A/view>

Nooruddin, Ifran. 2020. “Replication Data for: OAS Audit of Bolivia General Election 2019.” Harvard Dataverse, August 19.

<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/SGOFSC>

Organization of American States (OAS). 2019a. “Statement of the OAS Electoral Observation Mission in Bolivia.” October 21.

https://www.oas.org/en/media_center/press_release.asp?sCodigo=E-085/19

- _____. 2019b. “Electoral Integrity Analysis: General Elections in the Plurinational State of Bolivia: Preliminary Findings Report to the General Secretariat.” October 20. <http://www.oas.org/documents/eng/press/Electoral-Integrity-Analysis-Bolivia2019.pdf>
- _____. 2019c. “Electoral Integrity Analysis: General Elections in the Plurinational State of Bolivia: FINAL REPORT.” October 20. <http://www.oas.org/fpdb/press/Audit-Report-EN-vFINAL.pdf>
- _____. 2020a. “Press Release on Disinformation Campaign Regarding the Role of the OAS in the Bolivian Elections.” June 16. https://www.oas.org/en/media_center/press_release.asp?sCodigo=E-064/20
- Órgano Electoral Plurinacional (OEP). 2019. “Reglamento para las Elecciones Generales 2019: Aprobado Mediante Resolución TSE-RSP-ADM No. 230.” May 24. https://www.oep.org.bo/wp-content/uploads/2019/05/Reglamento_Elecciones_Generales_2019.pdf
- Rosnick, David. 2019. “Unnatural Claims in a ‘Natural Experiment’: Escobari and Hoover on the 2019 Bolivian Elections.” Washington, DC: Center for Economic and Policy Research, December 12. <https://cepr.net/report/publications-reports-unnatural-claims-in-a-natural-experiment-escobari-and-hoover-on-the-2019-bolivian-elections/>
- _____. 2020. “The Ends Don’t Justify the Means: Yet Another Analysis Fails to Support the OAS’ Delegitimization of Bolivian Elections.” Washington, DC: Center for Economic and Policy Research, May 27. <https://cepr.net/report/the-ends-dont-justify-the-means/>
- Tribunal Supremo Electoral (TSE) de Bolivia. 2020. Data obtained via formal email request. May 25.

Appendix: Escobari and Hoover Robustness Check

In our response to Escobari and Hoover, we pointed out multiple issues with their regression analysis.⁴⁴ One issue is that they incorrectly denominate the vote shares by the “Votos Válidos” entered (often incorrectly) on the tally sheets. This is just a check on the other numbers and has no official meaning. For purposes of computing the margins of victory in making a legal determination of the winner, the proper denominator is the sum of valid votes counted.

Another issue we raised was the use of unweighted observations. A tally sheet that broke 3 to 2 in favor of Mesa would have a margin of -20 percentage points and have the same importance as a tally sheet that broke 101 to 99 for Morales, for a +1-percentage point margin. Escobari and Hoover do claim in their paper that they weighted their regressions, but replications clearly suggest that they did not.

We successfully replicated (to within the published precision) every statistic of Escobari and Hoover’s Table 3, Columns 1–3, by employing the improper denominator and no weights. We see that with respect to the precinct-level analysis, the statistical significance of the results is not at all robust to our concerns. Once the regressions are weighted, the marginal significance of Escobari and Hoover’s variable of interest vanishes.

⁴⁴ The authors have since updated their paper with additional arguments based on non-public data. They have failed to respond to repeated requests to share the data or to answer questions about the additional results they presented. We therefore do not consider the additional arguments credible, and thus will not provide them with further examination at this time.

Table 35

Escobari and Hoover Results: Reported, and Replications

Number in parenthesis is the standard error of each statistic.

| | | Weight | Denominator | SHUTDOWN Coefficient | | |
|---------------------------|-----------------------|--------------------|-------------|----------------------|--------------------|--------------------|
| Controls | | | | None | Municipality | Precinct |
| As reported ⁴⁵ | | | | 16.26 [†] | 7.243 [†] | 0.365 [*] |
| | | | | (0.653) | (0.437) | (0.194) |
| Replications | None | Votos Válidos | | 16.26 [†] | 7.243 [†] | 0.377 [*] |
| | | | | (0.653) | (0.437) | (0.194) |
| | | Actual Valid Votes | | 16.27 [†] | 7.214 [†] | 0.360 [*] |
| | | | | (0.653) | (0.436) | (0.193) |
| | Votos Válidos | Votos Válidos | | 16.75 [†] | 7.789 [†] | 0.301 |
| | | | | (0.663) | (0.447) | (0.193) |
| | | Actual Valid Votes | | 16.75 [†] | 7.772 [†] | 0.288 |
| | | | | (0.663) | (0.447) | (0.192) |
| | Actual valid votes | Votos Válidos | | 16.79 [†] | 7.810 [†] | 0.307 |
| | | | | (0.663) | (0.447) | (0.193) |
| | | Actual Valid Votes | | 16.77 [†] | 7.766 [†] | 0.287 |
| | | | | (0.663) | (0.447) | (0.192) |

Sources and notes: TSE (2020), Escobari & Hoover (2019), and author's calculations.

*: statistical significance at a 10 percent level.

†: statistical significance at a 1 percent level.

The statistical significance of the SHUTDOWN variable is not convincing evidence of fraud; the variable serves as a catchall for anything unexplained with respect to the late votes. We would never expect that a model could possibly include every benign factor entering into the final election result (and Escobari and Hoover make no effort to do so). The fact that a slim but politically insignificant edge is explained by *something unidentified* which may be benign does not concern us. Once precinct controls are introduced, this simple model validates the original election result. The political significance is nil. Rather, the results reaffirm that an intense swing in favor of Morales should have been expected following the interruption in reporting.

However, to the extent that one might care about the statistical significance of this catchall, the result is neither robust nor present in the weighted regression model that Escobari and Hoover describe in their paper.

⁴⁵ The SHUTDOWN coefficient in column 3 appears to be misreported in Escobari and Hoover as all other numbers there agree within the reported precision, including its standard error, the statistics for the regression constant, the resulting number of observations, and the R²-value.

Annex: The OAS Statistical Analyses in the Final Report

The OAS statistical analyses in the Final Report (audit) of the 2019 presidential election in Bolivia are conceptually flawed. The heading for this section of the audit report, assembled by Irfan Nooruddin, reads: “FINDING 5: THE TREND OF THE FINAL 5% OF THE COUNT IS HIGHLY IMPROBABLE.” However, Nooruddin provides zero evidence suggesting that these tally sheets were in any way unlikely. Rather, Nooruddin simply observes the trends (or changes in trend) and simply declares that the results are “unusual.”

Nooruddin’s primary focuses on “discontinuity” in his computed trends in support of each of the two most successful parties in the election: incumbent Evo Morales, representing MAS, and former president Carlos Mesa, representing CC. Nooruddin argues that the last 5 percent of votes counted “favored Mr Morales substantially ... not just different than earlier parts of the [count] but also sharply different than the trend on the other side of the [95 percent] threshold.”⁴⁶

Indeed, if votes had been counted at random throughout the process, it would be alarming if the last 5 percent of the vote indicated sharply higher support for Morales when contrasted with the first 95 percent. Suppose Morales obtained 46 percent of the vote over 5.7 million votes selected at random, but 48 percent of the vote in the last 300,000. This would be possible, but incredibly unlikely — about one in a million.

However, in the Bolivian election votes were not counted at random; votes were grouped onto tally sheets. With only 34,555 tally sheets, rather than 6 million individual votes, the chance of the last votes randomly favoring Morales becomes much greater.

Worse still, the capital cities — where the opposition won by 18 percentage points — accounted for 45 percent of the first 95 percent of votes verified in the preliminary count. The remaining areas — where Morales won by 33 percentage points — counted for 66 percent of the last 5 percent of votes verified.

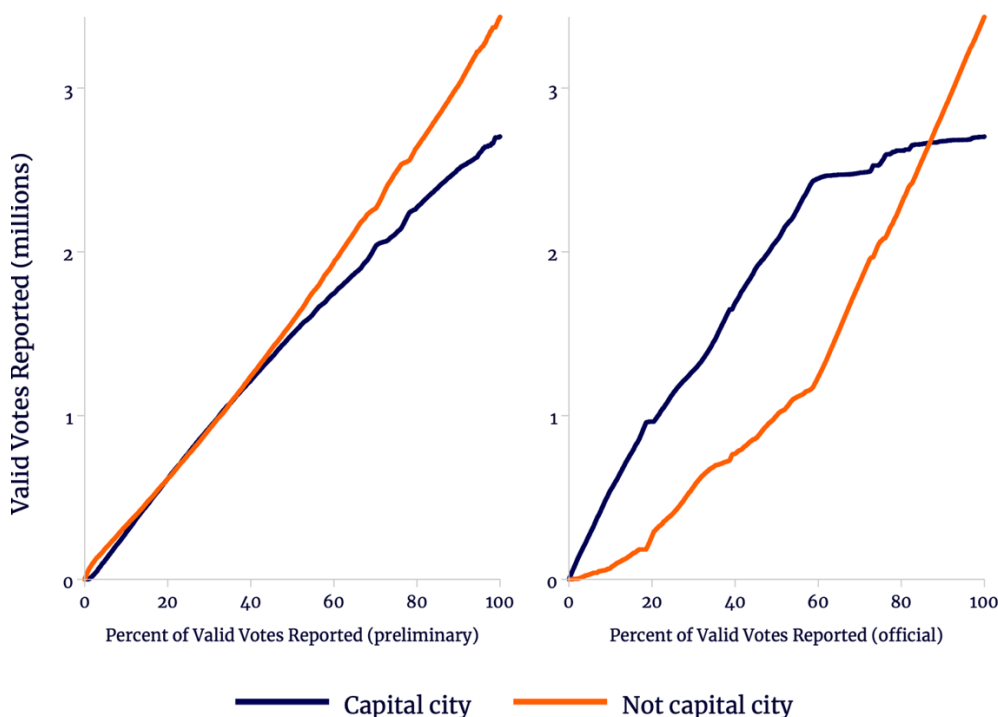
The official count was even more segregated; it required hand delivery of tally sheets to various department offices. Obviously, results from precincts nearest the offices arrived first,

⁴⁶ OAS (2019c).

while results from very remote and difficult-to-reach areas of Bolivia arrived last. Consequently, 93 percent of the last 5 percent of the official count came from outside the department capitals.

FIGURE 10

Progression of Vote Counts: Capital Cities and Other Areas



Source: TSE (2020) and author's calculations.

In both counts, the capital cities were counted disproportionately early when compared to outlying areas, and this had an impact on the support for each candidate over time. Of course, dividing the election between capital cities and everywhere else does not tell the whole story. Some capital cities (and parts of capital cities) favored Mesa more heavily than others; some outlying areas favored Morales more heavily than others. But the fact is, the ordering of the counts mean we ought to expect Morales to perform better on the late vote than he had previously.

In a very real sense, then, we simply care less whether or not there exists a discontinuity. Rather, the more important question is whether any discontinuity, or change in “trend,” or any other difference over the count appears unexpectedly.⁴⁷

⁴⁷ In any case, Idrobo, Kronick, and Rodríguez (2020) convincingly argue that Nooruddin’s observation is not of a feature of the data, but an artifact of peculiar choices in constructing his trendlines.

To answer the question of whether Nooruddin's observations were surprising or merely expected, we reconsider four figures from Nooruddin's analyses included in the OAS final report. First, we replicate the figures based on TSE data and Nooruddin's code. Second, we project the last 5 percent of the count based on the official results of the first 95. Third, we reanalyze the election by re-creating Nooruddin's figures based on the predictions.

We find that the projected results retain the same features that Nooruddin observes in the official data. The fact that these features are not unexpected based on the earlier data strongly suggests that the late results are not systematically irregular but that the tally sheets counted last, for whatever reason, happened to arrive from areas disproportionately favorable to Morales.

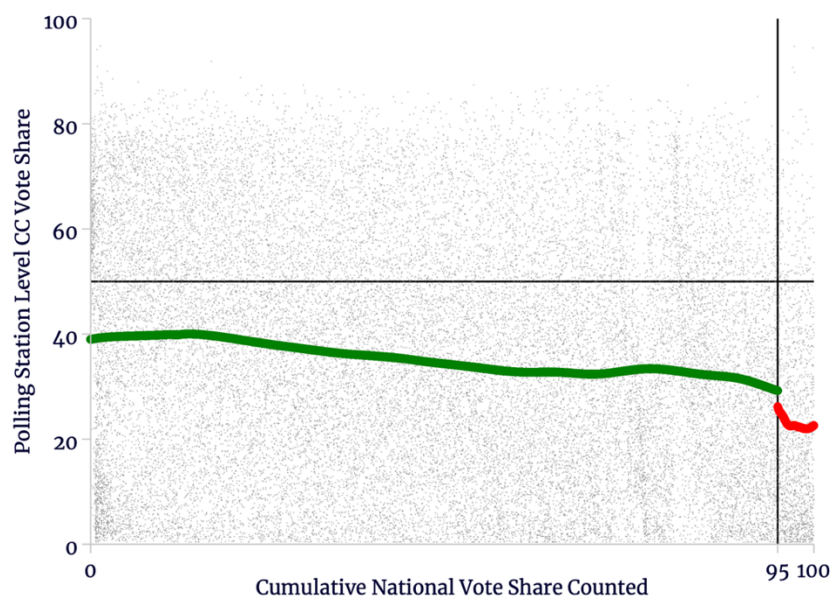
Along the way, we confirm several suspected data errors in Nooruddin's analyses.

Replication

Our first replication is the figure at the top of page 88 of the OAS final report. This intends to show, over the course of the count, the share of all ballots (including blank and null votes) cast in favor of Mesa, representing CC. Nooruddin includes two running-mean trendlines: one for the first 95 percent of ballots, and another for the last 5 percent.

FIGURE 11

Replication of OAS Page 88 (Top)



Source: TSE (2020) and author's calculations.

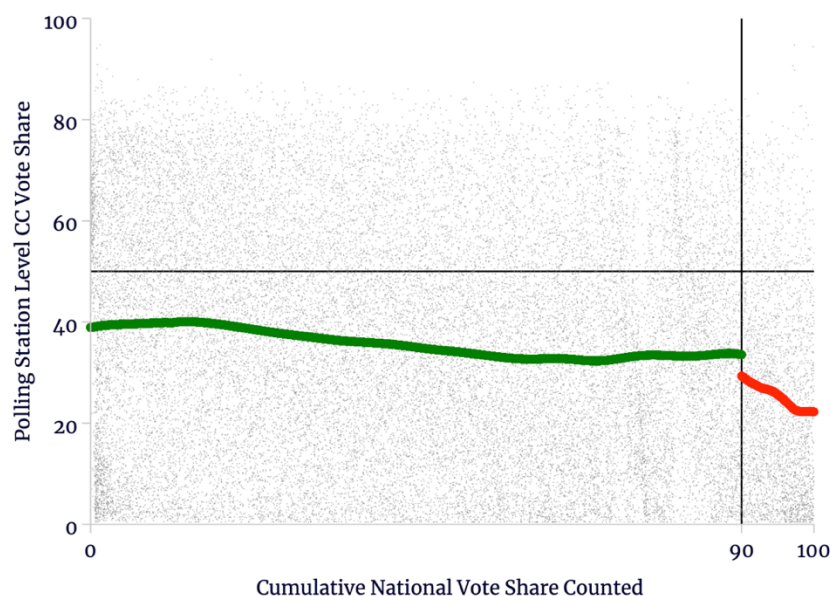
Many data concerns arose in reproducing **Figure 11**.

First, it is not clear how Nooruddin chose the 95 percent mark. In other studies, the focus has been on the last 16 percent of the count — the votes that came after the interruption. Though he describes the point “when approximately 90% of the vote had been counted” as “of immediate interest,” Nooruddin offers no explanation as to why.

Nooruddin argues that this “break in the CC’s vote trend is clearly visible ... and begs further scrutiny.” However, he also argues “we would not expect to see such a sharp discontinuity around an *arbitrary* point such as the 95% threshold used above.” [*Emphasis added.*] Indeed, if the 95 percent mark is indeed arbitrary and unusual, then we should observe no such break at 90 percent. This is not the case; there is nothing unique about the 95 percent mark.

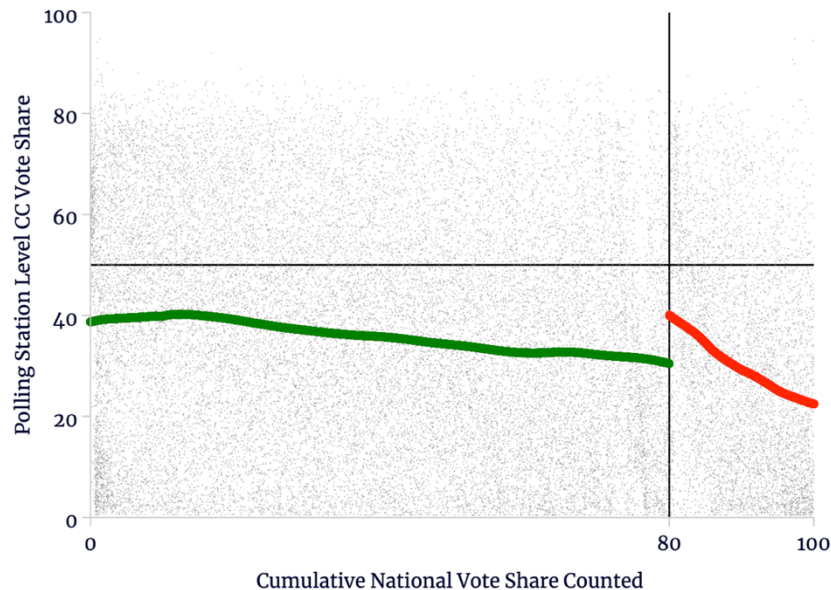
FIGURE 12

Mesa Share of *Emitidos*: Break at 90 Percent



Source: TSE (2020) and author's calculations.

At 80 percent, Nooruddin’s story reverses completely: rather than a sudden fall in the CC vote share, there is a sudden rise at 80 percent, followed by a gradual decline.

FIGURE 13**Mesa Share of *Emitidos*: Break at 80 Percent**

Source: TSE (2020) and author's calculations.

We do not believe any of these breaks are meaningful. Certainly, there was a swing in support away from Morales and toward Mesa around what here is presented as the 80 percent mark. As we discuss in our report “Observing the Observers,” it was about this time that a large number of tally sheets from opposition-heavy Santa Cruz were verified.

That is, these graphs do not depict sudden shifts in Mesa’s support, but shifts in the geographic composition of tally sheets. There is no reason to think the geographic composition of tally sheets would transition smoothly over time, so such shifts are apparent throughout the data. The individual trendlines so heavily smooth the data that they hide many other apparent shifts.

The significance of the break is therefore circular: Nooruddin’s 95 percent mark only becomes interesting because he chose 95 percent.

Second, the official data provided by the TSE leaves blanks in place of zeros where zeros are plainly indicated. For example, on tally sheet 81083, MAS received two votes, yet the CC total is blank. Nooruddin does not correct for this, so rather than recognizing that CC received 0 percent of the votes on the tally sheet, the tally sheet simply disappears from the figure.

Third, the votes are computed relative to *emitidos* — valid votes plus blank and null votes. It is important to recognize that the election is decided, legally, upon the valid votes alone. The votes and margins (differences between the top candidates) required to determine a victor are percentages of the valid votes, and not *emitidos*.

Fourth, Nooruddin's ordering of the tally sheets conflicts with the public data. Nooruddin considers a tally sheet "counted" as soon as it is timestamped as verified in the preliminary data, so long as the tally sheet was verified by the end of the preliminary count. Hundreds of verified tally sheets were set aside for further scrutiny. Some were not reported to the public until approved, and these Nooruddin considered "counted" at the time they were verified rather than at the time when they were approved. Others were never approved; inconsistently, Nooruddin considered these to be not yet counted, though already verified. Either the time of approval is important for sorting the tally sheets, or it is not.

Fifth, Nooruddin includes in his graph only the tally sheets verified and not awaiting approval in the preliminary count. Only 96 percent of *emitidos* in the election are represented in the graph; a sizable majority (79 percent) of the remaining *emitidos* in the election were on tally sheets verified but never approved as part of the preliminary count. Thus, tally sheets to the left of Nooruddin's "95" mark include only 91 percent of total *emitidos*.

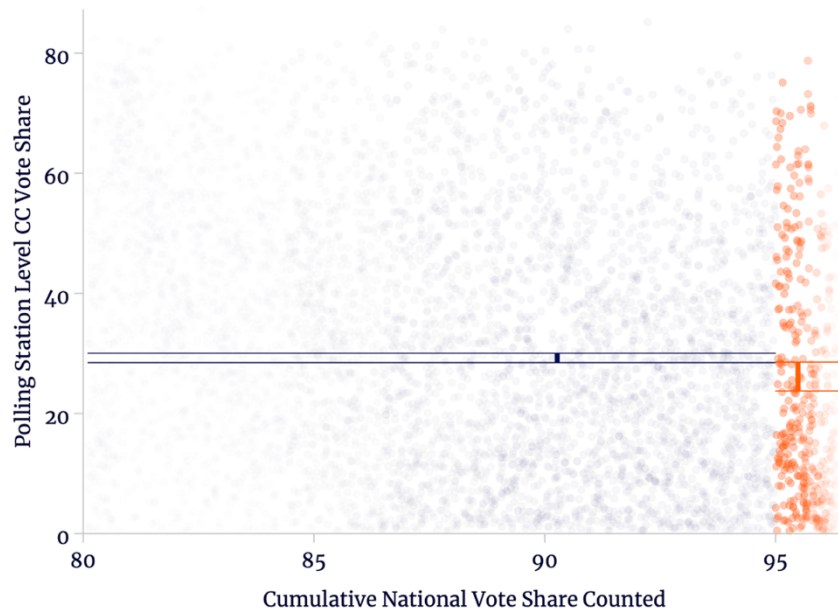
Sixth, Nooruddin does not explain his nondefault choices in estimating his trendlines. This is more fully explored in Idrobo, Kronick, and Rodríguez (2020). However, as an illustration we pick out one critical choice.

Nooruddin argues that the important feature is the difference in two trendlines at the 95 percent mark. Nooruddin's trendlines on either side of the 95 percent mark are simply weighted averages of Mesa vote shares on each side — the closer to the 95 percent mark, the more heavily weighted the tally sheet. However, Nooruddin selected disproportionate "bandwidths" for his calculations.

What do we mean by this? In Figure 5, we see the tally sheets used in the calculation of each trendline *seen at the 95 percent mark*. The darker the color, the more weight the tally sheet has in each computation. Though both calculations are meant to estimate Mesa's support at 95 percent, the tally sheets to the left are 10 times further away from 95 than the tally sheets to the right. In effect, what Nooruddin observes is not a change from 94.9986 to 95.0015 percent, but a change from 90 to 95.5.

FIGURE 14

Nooruddin's Estimate of Mesa Support Just Before and After 95 Percent



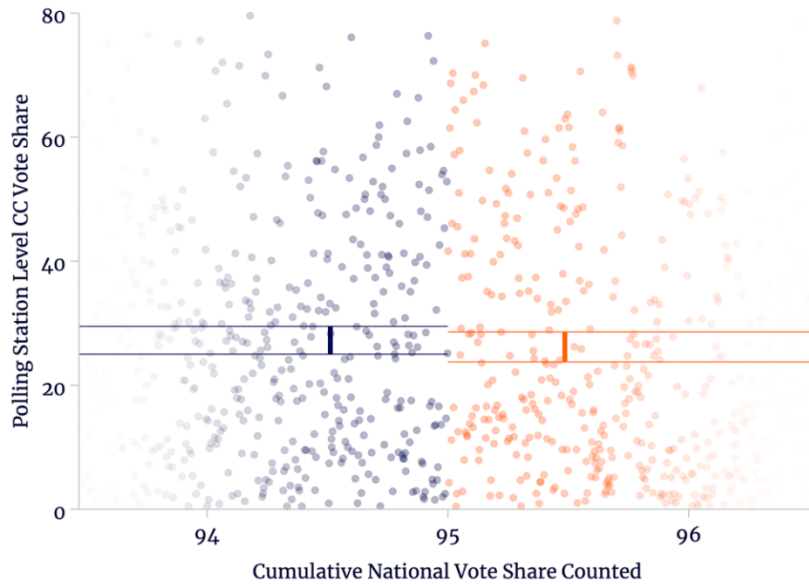
Source: TSE (2020), Nooruddin (2020), and author's calculations.

This poses a problem if Mesa's support has a steady downward trend. The trendline at 95 will underestimate support just to the right of 95 and overestimate support to the left by even more. This will look like a sudden fall in Mesa's support even if no such break exists. Idrobo, Kronick, and Rodríguez (2020), in their Figure 3(e) illustrate this effect on simulated data, and show that a slightly different approach to estimating the trend fixes that problem.

However, we may also improve the accuracy of the estimate by simply focusing more clearly on the tally sheets close to the 95 percent mark. If we equalize the number of tally sheets on each side of 95 percent, as in Figure 6, the apparent discontinuity vanishes.

FIGURE 15

Balanced Reestimate of Mesa Support Just Before and After 95 Percent

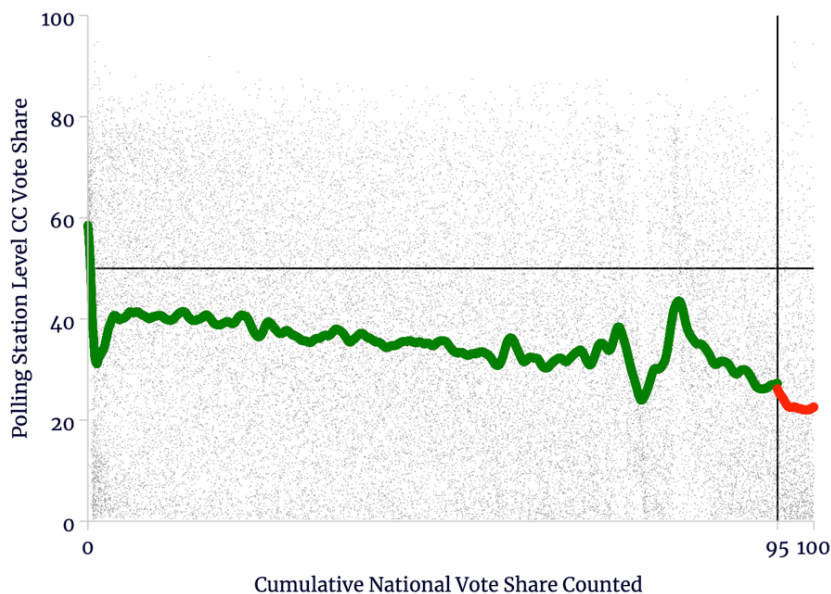


Source: TSE (2020), Nooruddin (2020), and author's calculations.

If we apply the smaller “bandwidth” to the entire trendline left of 95 percent, not only does the discontinuity vanish, but the “Santa Cruz effect” around 80 percent becomes visible, as does the Morales-favored “El Alto” shift that preceded it. It is clear that Mesa’s support largely trends downward through most of the count with nothing notable at the 95 percent mark.

FIGURE 16

Reanalysis with Equalized Bandwidths



Source: TSE (2020), Nooruddin (2020), and author's calculations.

We therefore affirm Idrobo, Kronick, and Rodríguez's (2020) finding that Nooruddin's jump was an artifact of his choices, rather than a clear feature of the election data.

Prediction and Reanalysis

Let us put aside all these concerns regarding Nooruddin's methods and for the purposes of argument adopt exactly his approach. Because shifts in vote shares may be driven by changes in which geographical areas are counted and because these changes need not be smooth, the question is not whether a break in trend is observed but whether a break in trend is predictable based on the information available.

To answer this question, based only on the first 95 percent of the count we predict voting on the last 5 percent of tally sheets (per Nooruddin's definition) and apply his methods — including unofficial denominator, dropping zeros, use of local means, and nondefault bandwidths — observing the resulting predicted trend.

To make our predictions, we follow the procedure described in the Technical Appendix of our “What Happened in Bolivia's 2019 Vote Count?” with the additional consideration of the urban/rural divide as defined by Chumacero. The procedure is roughly as follows:

Suppose that most tally sheets from a particular precinct are among the first 95 percent, but one is not yet counted. We might randomly select any of the sheets already counted as representative of the precinct. If Mesa received 40 percent of the vote on the randomly selected representative, then we predict that Mesa receives 40 percent of the vote on the uncounted tally sheet.

If we suppose there were two uncounted tally sheets, then we separately select a representative for the second uncounted tally sheet, and so on.

If the precinct is not counted at all prior to the 95 percent mark, however, the procedure becomes more complicated. Perhaps the locality contains two other precincts, both of which had tally sheets counted among the first 95 percent. It may be that our uncounted precinct looks like one, or both, or neither of the other two. Not knowing which, we randomly select a representative precinct, and draw all representative tally sheets from that particular precinct.

This doesn't necessarily solve our problem. If there are no precincts from the locality that had been at least partially counted, we must turn to the urban or rural part of the corresponding municipality. From the urban or rural part of the municipality, we may be able to randomly choose a representative locality, and from that locality choose at random representative precincts, and then representative tally sheets.

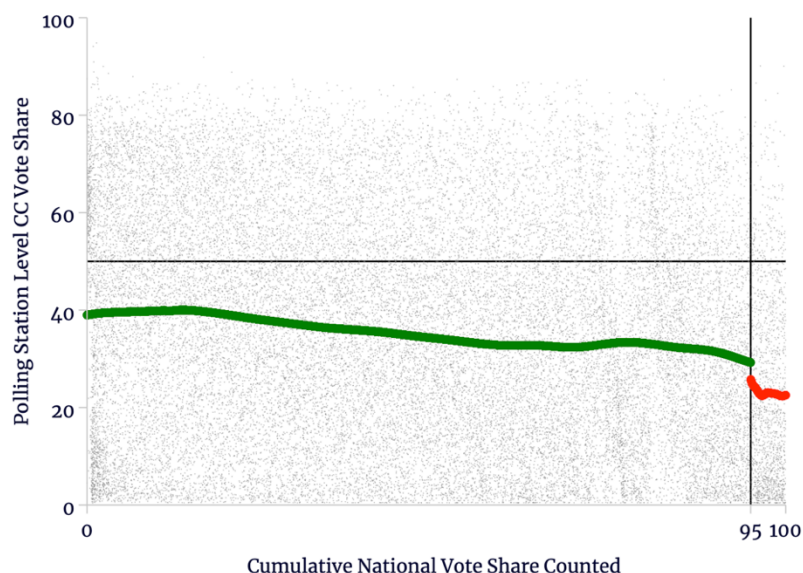
This process eventually works its way out. If there were, for example, no tally sheets from Colombia, we may randomly choose between urban or rural Bolivia as representative. We may then pick from the urban (or rural) parts of each department, province, and so on.

Once we have a prediction for every uncounted tally sheet, we may repeat Nooruddin's analysis with our predictions in place of the official data. Because the representative geographies may or may not accurately represent the uncounted geographies, this introduces an extra level of uncertainty in our predictions. Some representatives will underestimate Mesa's support; others will overestimate it, but not all errors will be equal. Some random choices will have larger effects on the prediction than others.

We might then see a figure very much like — but not identical to — Nooruddin's original. Though we again see a difference in trendlines at the 95 percent mark, this is what we expect based on the pre-95 data. Again, we believe this is a consequence of Nooruddin's analytical choices, but given those choices, Nooruddin's apparent discontinuity is entirely expected.

FIGURE 17

Reanalysis with Projected Post-95 Tally Sheets



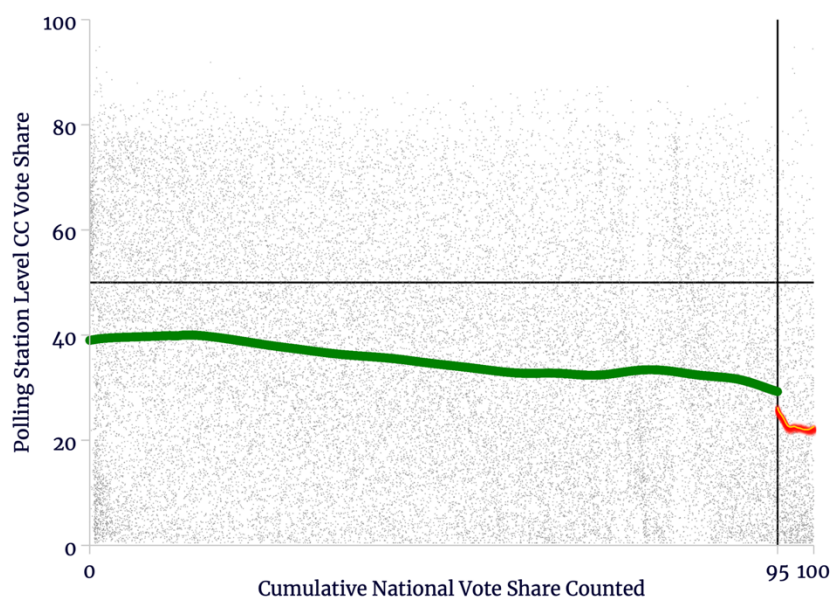
Source: TSE (2020), Nooruddin (2020), and author's calculations.

Having made predictions for all the uncounted tally sheets and performed Nooruddin's analysis on the expected data, we may reselect at random all the representative geographies. Some representatives that underestimated Mesa's support may now overestimate it, and vice versa. Repeating the analysis many times produces a range of plausible estimates of Mesa's support in the last 5 percent.

In the figure below, we summarize 150 such projections in red. The thin yellow line shows Nooruddin's result, indicating the shift in Mesa's support is completely predictable. Both the fall immediately after the 95 percent mark and the "L" shape thereafter are plainly visible in the simulated results. Nooruddin does not explain why he thinks the result is suspicious, but it appears that — given the order in which the tally sheets were counted — that the late results were not unexpected.

FIGURE 18

Reanalysis with 150 Projections of Post-95 Tally Sheets



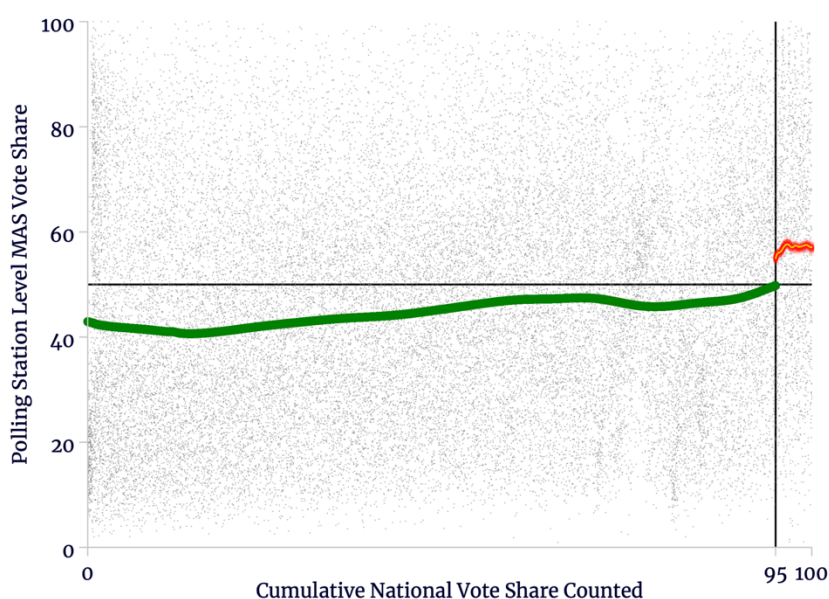
Source: TSE (2020), Nooruddin (2020), and author's calculations.

Replication, Prediction, and Reanalysis of Other Figures

Just as we are able to replicate and reanalyze Nooruddin's figure describing Mesa's vote share during the preliminary count, we may reproduce the corresponding graph for Morales's support described at the bottom of page 88 of the OAS final report. All the previous data concerns again apply here, but with those in mind the figure is reproducible. And again, the 150 projections indicate that the official results fall well within a relatively narrow range of outcomes based on the first 95 percent of tally sheets.

FIGURE 19

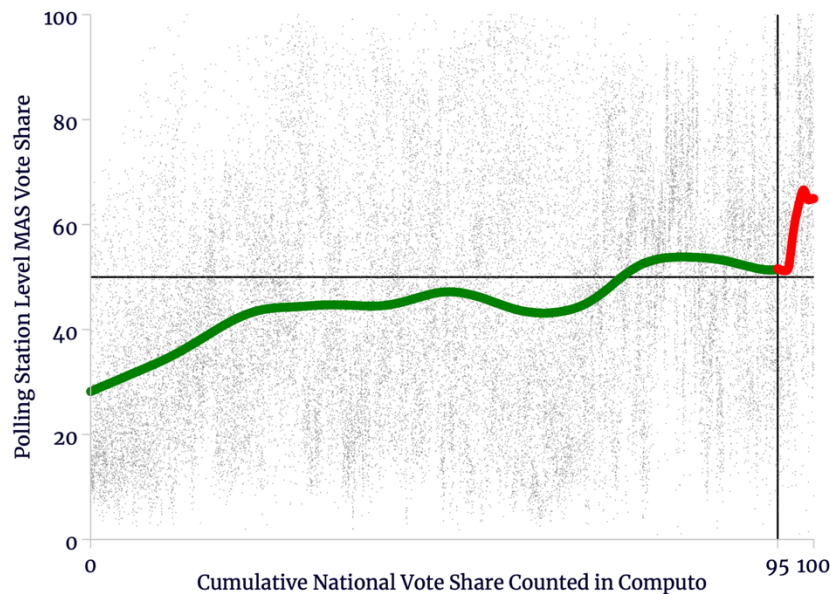
Replication and Reanalysis of OAS Page 88 (Bottom)



Source: TSE (2020), Nooruddin (2020), and author's calculations.

While trial and error allowed researchers to detect Nooruddin's unusual data choices and thereby replicate reasonably closely the graphs for the preliminary count on page 88, the same could not be said for the corresponding graphs on page 92 describing the official count. However, on August 19 — more than nine months after the publication of the report — Nooruddin finally responded to public pressure generated by press coverage and shared his data. We quickly discovered that while Nooruddin had ordered his tally sheets by the date of second transcription in the official count, he had done so alphanumerically rather than chronologically.

Accounting for this missorting allowed reproduction of the graphs in the report. First, we see Nooruddin's progression of Morales's share over the official count.

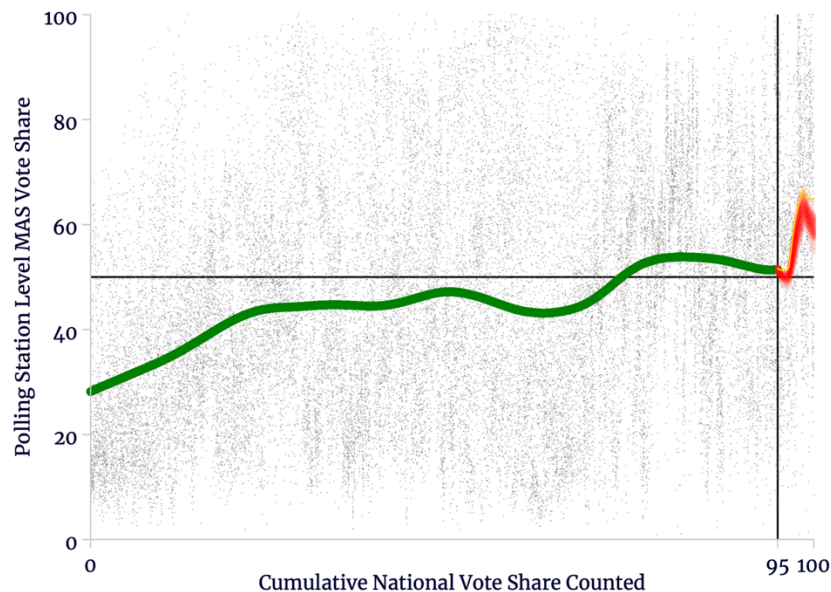
FIGURE 20**Replication of OAS Page 92 (Top)**

Source: TSE (2020), Nooruddin (2020), and author's calculations.

This appears to have very slight differences when compared to the figure at the top of page 92. For example, the trendlines do not quite touch the 50 mark around 95 percent complete. However, as this feature is generated in a graph produced with Nooruddin's own code, we accept our figure as exactly replicating Nooruddin's.

Many of Nooruddin's errors from analysis of the preliminary count carry over to the official. The reason for dividing at 95 percent is not explained. Tally sheets where Morales received zero votes are dropped. The choice of denominator is unofficial, and Nooruddin does not explain his nondefault choices.

As discussed above, the tally sheets are completely missorted to the point that those in the "last 5 percent" do not in any meaningful sense correspond to the last 5 percent of the official count. However, if we accept his chosen timing of the tally sheets, the late swing toward Morales is far from surprising.

FIGURE 21**Reanalysis of OAS Page 92 (Top)**

Source: TSE (2020), Nooruddin (2020), and author's calculations.

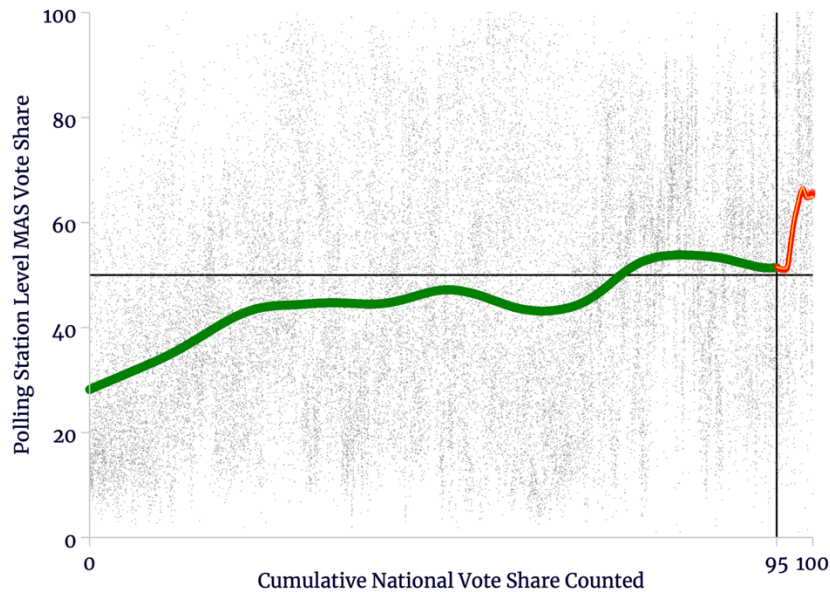
To be sure, the official data lies on the upper range of our simulations. This is in part due to the stronger geographic ordering in the official count. Because the opposition-favored areas were overwhelmingly counted early, it is disproportionately less likely that any representative geography favors Morales. Regardless, there is little question that Nooruddin should have expected to see a sharp swing in his trendline.

There are at least two ways that we can improve upon these estimates. First, we could use prior election data or socioeconomic status so that our randomizer tends to select better representatives. Second, we may take note of the fact that by the time the official count reached 95 percent, we had information from the preliminary count. If we had already observed that the remaining tally sheets in particular had strongly supported Morales, there is no reason we should be surprised that the late votes in the official count broke in his favor.

In the graph that follows, we project only those tally sheets in the last 5 percent of the official count that had not been reported preliminarily by 19:40:57 on election night.

FIGURE 22

Reanalysis of OAS Page 92 (Top) with More Complete Information



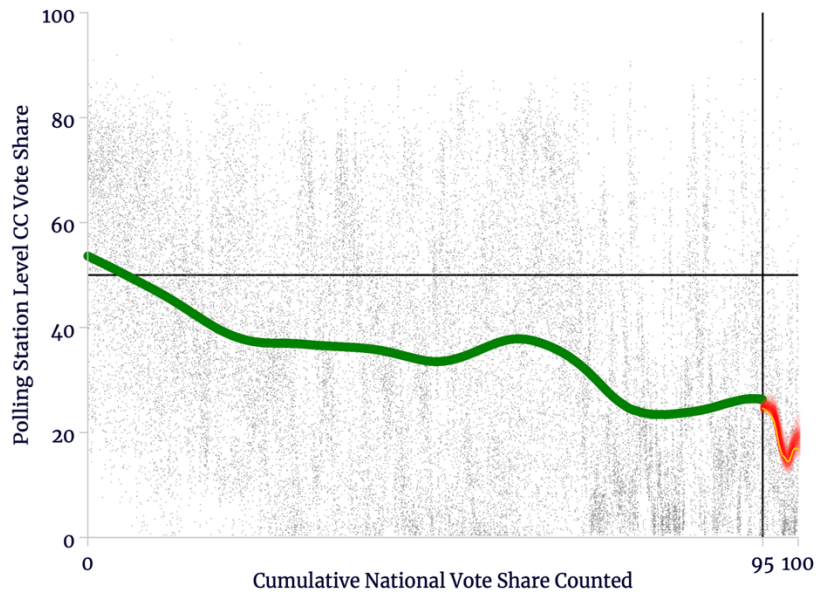
Source: TSE (2020), Nooruddin (2020), and author's calculations.

Even discarding preliminary count data that arrived after the interruption in reporting, there is so much information available by the time the official count reaches Nooruddin's 95 percent mark that the swing is very accurately and precisely predicted.

Likewise, we successfully reproduce Nooruddin's figure at the bottom of page 92. Using only early data from the official count, predictions are a bit favorable to Mesa.

FIGURE 23

Reanalysis of OAS Page 92 (Bottom)

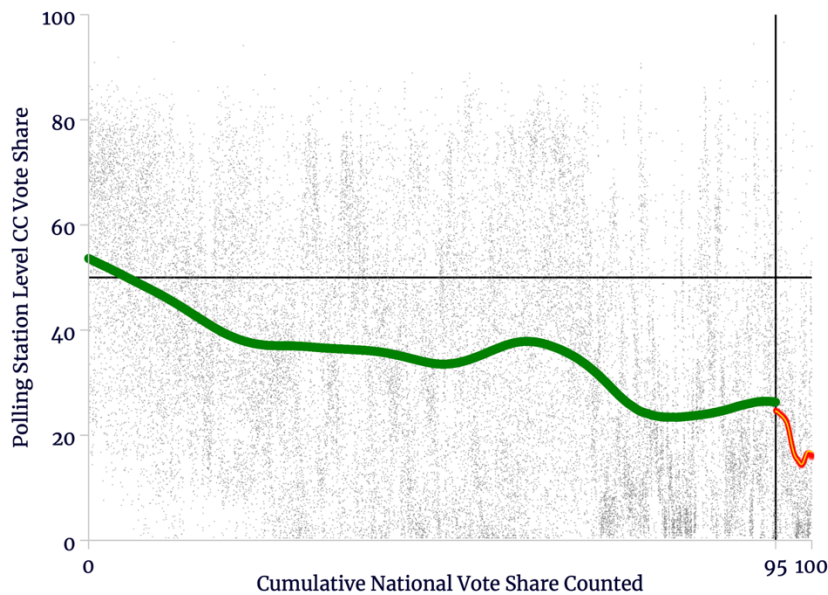


Source: TSE (2020), Nooruddin (2020), and author's calculations.

However, using early preliminary count data as well, we find that the late swing is completely expected.

FIGURE 24

Reanalysis of OAS Page 92 (Bottom) with More Complete Information



Source: TSE (2020), Nooruddin (2020), and author's calculations.

Conclusion

Even taking as given Nooruddin's erroneous ordering of the tally sheets and odd methodological choices, the late breaks in support for Morales are predictable with considerable accuracy. Despite the OAS's characterization of party shares late in each count as "highly improbable" and "extremely unusual," Nooruddin merely asserts the shifts self-evidently "cannot be explained away" and their existence "strains credulity." However, these shifts are clearly explained by taking into account the origin of the remaining tally sheets.