Nickels Before Dimes

Misleading Design in Escobari and Hoover's "Natural Experiment" and Bolivia's 2019 Election

By David Rosnick and Aileen Wu*

November 2022

CEPR

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 (CEPR). Aileen Wu is an intern at CEPR.

Contents

Acknowledgements

The authors thank Dan Beeton, Brett Heinz, and Jake Johnston.

Executive Summary

In October 2019, the Plurinational State of Bolivia held legislative and presidential elections. To win in the first-round vote, the leading presidential candidate would need to win a majority of the valid votes cast, or 40 percent with a 10 percentage point lead. Because of division within the opposition, surveys suggested that support for the incumbent, Evo Morales, lay close to the latter threshold.

With 84 percent of a rapid, preliminary count of the election results completed, the Tribunal Supremo Electoral (TSE) announced partial results. At that point, Morales led the next-closest candidate, Carlos Mesa, by 7.9 percentage points. However, the outstanding polling stations were expected to strongly favor the incumbent. A Morales victory in the first round was predictable based on the data available at the time of the announcement. Indeed, the final, official result showed Morales with a sufficient 10.56 percentage point lead.

However, the Electoral Observation Mission (EOM) of the Organization of American States (OAS) cried foul, declaring the updated results represented "an inexplicable change in trend that drastically modifies the fate of the election."[1] This lent international institutional legitimacy to opposition protests disputing the results. These protests turned violent, threatening the lives of high-level government officials and their families alike. And following a mutiny of police forces and a public call for Morales's resignation by the head of the military, Morales and his vice president resigned and fled to Mexico, where the Mexican foreign ministry had granted Morales asylum given "the emergency situation that he faces in Bolivia, where his life and safety are at risk."[2] Shortly afterward, an extra-constitutional government headed by Jeanine Áñez was installed and Bolivian security forces acting under the de facto government's direction violently repressed Indigenous protesters and perpetrated massacres at Senkata and Sacaba.

Since the OAS statement, several researchers have attempted to offer statistical evidence for the OAS claims. In October 2020, just days before new elections, Diego Escobari and Gary Hoover released an extensive paper arguing that they had identified fraud sufficient to change the outcome of the 2019 election.

None of Escobari and Hoover's claims hold up to scrutiny. In one sense, this could be an issue of semantics. If, in choosing to work from overly simplistic models, Escobari and Hoover inferred

1 OAS (2019c).
2 Gobierno de México (2019).

merely that their models do not fully explain the trend of increased support for Morales in the later-reported election results, then their conclusions would be less controversial. In labeling this modeling failure "fraud," they implicitly assert that they have exhausted all benign explanations. However, they make no actual claim of exhaustion.

Only at first do they employ a "natural experiment" approach to analyzing fraud. There, they assume that the partial results presented by the TSE the evening of the election were free of fraud, while the remaining, uncounted polling stations are potentially contaminated. However, the extent of fraud they claim to measure in these late polling stations is implausible. The official results in precincts counted entirely after the announcement would need to overstate Morales's support by 84 percentage points for their conclusions to hold.

Escobari and Hoover later abandon the claim to a natural experiment and extend their search for fraud to every single polling station in the election by contrasting results with a nationwide referendum from 2016. However, the models they apply are not well suited for the task of measuring election fraud.

Rather, when applied to simple fraud-free synthetic election data, their methods fail, reliably detecting fraud where none exists. The natural conclusions are that their approaches to fraud detection are faulty and that their claims to have measured fraud in the 2019 Bolivian elections are not supported.

Introduction

One month after the 2019 presidential election in Bolivia, Diego Escobari and Gary Hoover released a paper alleging that the election was fraudulent — that incumbent president Evo Morales failed to garner votes in sufficient numbers to win in the first round of the election.[3] Their argument was that later-counted tally sheets disproportionately favored Morales and that absent that inexplicable change the election had been headed toward a second round. Nearly a year later, they greatly expanded this paper to make wider claims of statistical evidence of fraud.

3 Escobari and Hoover (2019). It is worth noting that as of August 2022, Escobari and Hoover (2019) is no longer available on the SSRN repository, although to the authors' knowledge, no formal retraction has been issued. It is referenced in the present exercise for the purpose of explaining as fully as possible the varied approaches to this line of inquiry.

It is important to recognize that a claim of fraud rooted in statistics requires very strong modeling assumptions. When evidence is physical — such as a parallel set of financial accounts — the case for fraud is less complicated. When the evidence is that there remains an unexplained bias otherwise unaccounted for within a statistical model, labeling this bias "fraud" implicitly rules out any and all benign explanations. To justify this assumption can be a monumental task. How can a researcher incorporate every conceivable benign factor into an election model? Even past elections may be of limited use in this regard. If a candidate's vote total is low in some seemingly random town, has the researcher accounted for a candidate failing to visit that town — or worse — mispronouncing the name of the popular mayor, or refusing the local delicacy? Perhaps the candidate garnered more votes than expected because they promised to connect the town with high-speed rail. The number of benign factors is effectively without limit, and no researcher can account for them all.

Sometimes, having a large number of small, unincorporated factors does not cause much trouble. If a large number of pluses and minuses are scattered uniformly throughout a vote count, they will not cause the later-counted votes to look very different than those counted earlier. However, every benign reason why late-counted votes may differ systematically from early-counted votes must be taken into account or be mistaken for fraud.

To illustrate, consider this tale of two roommates.

The Parable of Nickels Before Dimes

Alex reminds Blake, "Don't forget it's my night to pick."

"We're a dollar thirty short," says Blake. "We can't afford your fancy pizza. Let's just go get those awesome sandwiches from the corner store."

"A dollar thirty?" responds Alex, "You really don't want pizza, do you? Go check the swear jar. There's enough there."

"Alex, there's only 20 coins in here. That's not enough." Blake starts pulling out coins, one by one. Ten coins in, they hold eight nickels and two dimes. "Sorry, Alex, I'm halfway done counting but so far we have only 60 cents. I'm going downstairs for sandwiches."

By the time Blake returns with their sandwiches, Alex is livid. "I counted the rest of the jar. A dollar thirty more for pizza, and we have a dollar fifty. I told you there was enough money. I can't believe you'd lie rather than get pizza."

"I can't believe you'd accuse me of lying, Alex. The first half of the jar came to 60 cents. That's only a dollar twenty total. I don't know where you got the extra 30 cents, but it wasn't from the jar."

The phone rings, and Alex picks up. It's their friend Charlie calling about meeting up this coming weekend. Alex, exasperated, tells the story of what happened. Charlie frowns. "Y'all will be much happier if you sort out some issues with your dinner decisions. But let's think this through. Just because the first ten coins were eight nickels and two dimes doesn't mean the other 10 counts were eight nickels and two dimes. Even if Blake had drawn at random from a jar of ten nickels and ten dimes, there is a 1 in 87 chance that eight of the first ten would be nickels."

"See!" Blake retorts, "That's not very likely, so Alex added money to the jar."

Charlie takes a deep breath. "One in 87 is not very likely, but not at all impossible. It could just have been bad luck. Yet there is a more innocuous explanation. Nickels are bigger than dimes. If Blake simply drew the first coin they touched, it would more likely be a nickel. Depending on how Blake happened to draw the coins the odds could have been one in eight, or maybe one in three. They may not have realized the bias in selecting nickels before dimes.

"It doesn't matter if Blake drew nickels before dimes deliberately or unwittingly. Deciding halfway through the count that there wasn't enough money and buying sandwiches was premature."[4]

Likewise, proposed theories of fraud must take particular care to account for "nickels before dimes." That is, fraud analysts must take care to not make unwarranted assumptions about the data.

In this manner, it is not unusual for candidates to receive unusually strong support late in the vote count because ballots are not counted at random. In the 2020 US presidential election, several states delayed counting mail-in ballots. As the mail-ins were expected to favor Biden,

4 In lieu of nickels and dimes, our Bolivian friends likewise may consider the larger, heavier 2-boliviano coin and the smaller, lighter coin worth 5.

Trump's lead predictably diminished and even disappeared as those ballots were eventually counted. Meanwhile, North Carolina counted mail-ins early, giving Biden a lead that reversed as in-person votes were fully counted.[5] After trailing all night, John Bel Edwards won reelection as governor of Louisiana in 2019 based on late reporting from Orleans Parish. Orleans is the largest parish in the state, accounting for nearly 8.5 percent of all votes cast. Edwards won by 80 percentage points there, having lost by 4.5 in the rest of the state.[6] In both cases, accounting for "nickels before dimes" helps discount the hypothesis that these "trend shifts" were the product of fraud.

What About Bolivia?

In the case of Bolivia's 2019 elections, suspicion of fraud centered on the public reporting of the preliminary results. The preliminary count, referred to as the TREP (for the Transmisión de Resultados Electorales Preliminares — the system of gathering, counting, and reporting these results) was unofficial and had been intended only to provide to the public a general idea of how the election had played out. The TREP was not at all critical to the election. In fact, with little complaint, the preliminary count was scrapped entirely in 2020.

Here is how the process worked. A random selection of eligible voters within each polling station were summoned for electoral jury duty. These citizens were called upon to administer their polling stations and to tabulate the results at the end of the day, filling out an *acta* — an official record of the vote totals.[7]

Tabulation errors were expected, and so the acta included checks to assist the jury in its duties. Nevertheless, ordinary citizens chosen at random are not always experienced in such matters,[8] and clerical errors were prevalent. One of the more frequent errors was to include blank and null

5 New York Times (2020), North Carolina State Board of Elections (2020), and North Carolina State Board of Elections (2022).

6 Louisiana Department of State (2019).

7 Actas are frequently referred to as "tally sheets," as they contain the totals. However, there were separate sheets containing actual tally marks for the purpose of counting the vote. In this paper, we refer to actas and tally sheets interchangeably.

8 According to the Electoral Observation Mission (EOM) of the Organization of American States (OAS):

> the EOM observed that the level of poll workers' training was in most cases insufficient, forcing them to constantly turn to the election notary in order to carry out their duties. This became more acute at the time the polls were closing, during the vote tallying, and when the tally sheets were being filled out. OAS (2019b), 16.

Note that this is the primary argument against the audit report vis-à-vis handwriting analysis as evidence of forged tally sheets. It is a clear leap of logic from overreliance on notaries to fraud.

votes in the space allotted for checking the total valid votes. Sometimes, an error would be caught and, per instructions, the error could not be repaired directly but a supplementary note — an *observación* — was placed on the acta. For instance, votes for one party might be recorded or duplicated in a field reserved for another party. This would not only trigger an observación if caught, but might result in an apparent mismatch between the party votes reported and the total valid votes declared. In some cases, jurors failed entirely to record the total valid votes.

Regardless, TREP operators in each precinct (one per ten polling stations) would then transcribe the results (using a phone app), take a photo of the acta, and submit these electronically to the Servicio de Registro Cívico (the Civil Service Registry, SERECI) in La Paz. Carbon copies of the acta would be distributed to the TREP operator, an electoral notary at the precinct,[9] the president of the jury, and any party delegates present. The original would be packed up with unclaimed copies and supporting material. The notaries would physically deliver these packages to the regional Tribunal Electoral Departamental (TED) offices. The original actas would then be opened, scrutinized, and reported in the official count known as the *cómputo*.

Meanwhile, SERECI workers would be assigned actas to verify drawn randomly from the pool of available image/transcription pairs. For much of the time, these workers were able to keep pace with the flow of data and so actas were initially verified in approximately the same order in which they had arrived. This was less true at the peak of transmission as they fell behind and the pool of available image/transcription pairs grew. Some actas were flagged for further scrutiny, i.e., for questions about observaciones or illegibility, or because transcription of the numbers did not match the transcription submitted electronically at the polling station. These actas were put aside in the TREP and many, though not all, were eventually included in the preliminary total.[10] Despite these checks, accuracy and completeness were sacrificed for speed in the TREP, in contrast with the slow and deliberate pace of the cómputo that the TSE would not complete for another five days.

9 There were 8,136 electoral notaries for 5,301 precincts. OEP (2019a).

10 It is worth noting that the OAS, in its audit of the election, would spin this process, saying the TSE "deliberately hid" actas in the system. OAS (2019a), 4. It appears that procedure was followed in not publicly reporting actas with conflicting transcriptions, unscrutinized observaciones, or that had not yet been assigned for transcription. It is not at all clear what the motive would be for publishing clearly uncertain results, let alone for "hiding" them. Even the OAS, in its final report, recommended that operators not be allowed to approve such actas. OAS (2019b), 30. Rather, the OAS claim relies on the public results at the time of the TSE announcement, with 29,526 of 33,048 (Bolivia-only) actas listed as transmitted, and only 27,782 actas verified. To some extent, the difference may be due to a backlog on the part of the SERECI workers. More likely, the majority of these actas were counted as transmitted for having sent in a transcription, though not in the verification pool for lack of a transmitted image to pair. Regardless, the 2019 TREP was not the first to see verifications lag transmissions in this way. In 2017, for example, with 89.34 percent transmitted, only 83.52 percent had been verified. ANF (2017). The 2016 process was a little better at the same stage, with 85.0 percent transmitted and 83.3 percent verified. La Razón Digital (2016).

At 19:50 on election day (October 20, 2019), the TSE held a press conference to report that with 83.9 percent of polling stations included in the preliminary count, Morales led runner-up Carlos Mesa by 7.87 percent of the valid votes.[11] In some precincts, all the polling stations were included in the announcement of partial results. We call these precincts "early" for having been fully included in the TSE announcement. Morales (tentatively) won these early precincts by 9.44 percentage points. Other precincts we call "split" with some "early" polling stations considered in the announcement, and other "late" polling stations still outstanding.[12] Morales led the early polling stations in split precincts by only 6.82.

Table 1

Preliminary Count Results At Time of TSE Announcement (19:40:57)

	Morales	Mesa	Difference	Valid Votes	Margin
	(a)	(b)	(c)=(a)-(b)	(d)	(e)=100(c)/(d)
Early precincts	971,216	776,098	195,118	2,066,601	9.44
Split precincts	1,384,428	1,173,902	210,526	3,086,378	6.82
Total	2,355,644	1,950,000	405,644	5,152,979	7.87

Sources: OEP (2020c) and authors' calculations.

To claim a victory in the first round, Morales would have to obtain a 10 percentage point margin of victory to avoid a run-off election.[13] This would require a 21 percentage point margin of victory at the remaining, uncounted polling stations — a little above 13 points more than the margin reported in the early results.

Such a swing had precedent in a failed 2016 referendum to eliminate Bolivia's presidential term limits. After 83.1 percent of the referendum results had been verified in that preliminary count, the "*Sí*" vote trailed by 7.6 percentage points.[14] By the final count, the referendum failed by only 2.6[15] — a 5 point swing. A swing in 2019 of half that size would be more than enough to hand Morales a first-round victory.[16]

11 The 2019 TREP reports are no longer available on the website of the Órgano Electoral Plurinacional (OEP), but are available at the Internet Archive (Órgano Electoral Plurinacional (Bolivia) [2020a]), along with the available acta images. This data may be found at the Internet Archive (Órgano Electoral Plurinacional (Bolivia) [2020c]); direct download is very slow.

12 A full list of polling stations, including their precinct assignments, may be found in the final, official, results archived at the Internet Archive (Órgano Electoral Plurinacional (Bolivia) [2020b]), but is also currently available at OEP 2022.

13 For a candidate to win in the first round, they must win a majority of the valid votes, or defeat the runner-up by at least 10 percentage points. For an outright majority, Morales would need almost 73 percent of the outstanding vote.

14 Opinión (2016).

15 OEP (2016a).

16 In 2016, the swing required net 23.8 percentage points support for the referendum over the last 16.9 percent of the count, compared to -7.6 percentage points over the first 83.1. Thus, the last part of the count favored the referendum by 31.4 percentage points, relative to the rest. Extending to 2019, half the difference of 31.4 percentage

It should be noted that the OAS sent a mission to observe the 2016 referendum and despite the late swing, the mission did not express concern. Rather, it does not require much of a skew in a sample to produce a significant swing in support. As an illustration, suppose we have 200 polling stations, of which 100 came from precincts that opposed the referendum 3:1 (a -50-percentage point margin in favor), and 100 from precincts that supported the referendum by 3:1 (a +50-percentage point margin in favor). Overall, these 200 polling stations neither favor nor oppose the referendum. If we take 85 of the former and 75 of the latter (a 53/47 split) however, this sample of polling stations would oppose the referendum by 3.1 percentage points. The rest would favor the referendum by 12.5. The difference in support between the sample of 160 stations and the remaining 40 stations is due entirely to a mild bias in the sampling of stations. In the earlier coin counting example, the average value of coins appeared to rise by three cents due to a bias in selecting *nickels before dimes*. Polling station selection bias can cause a swing in support in the same manner.

While it is not possible to match exactly the polling stations, or even the precincts, across elections, it is possible to make an imperfect comparison of precincts between 2016 and 2019. As a proxy, it is certainly imperfect. Nevertheless, we source historical data from several official websites as well as data provided via email directly from the TSE.[17] The TSE provided timestamps and results for both the TREP and cómputo. Within minimal timing differences, these match the public Excel spreadsheets that were generated in real time during the count. To help match precincts we also use the *padrón* (electoral registry) for each vote — in 2016[18] and 2019.[19] The polling-station level results in 2016 are available at Bolivia's Atlas Electoral.[20] For simplicity, we employed Rómulo Chumacero's definitions of urban vs. rural localities.[21]

Combining data from 2016 and the early results from 2019,[22] we can see that voting patterns between the two elections are not unrelated. In **Figure 1**, we see results in the four largest localities. These four account for nearly 40 percent of the entire electorate. Overall, the voters of El Alto favored both the referendum and Morales much more than did the voters of Cochabamba,

points is 15.7. Adding in the 7.87 percentage points seen over the 83.9 percent counted comes to 23.6 percentage points over the last 16.1 percent of the count. Morales's net support rising from 7.87 to 21 percentage points for an overall 10.4 percentage points would seem well within the bounds of historical plausibility.

17 Rosnick (2020a). All data and code are available at Rosnick (n.d.).

18 OEP (2017).

19 OEP (2019b) and OEP (2019c).

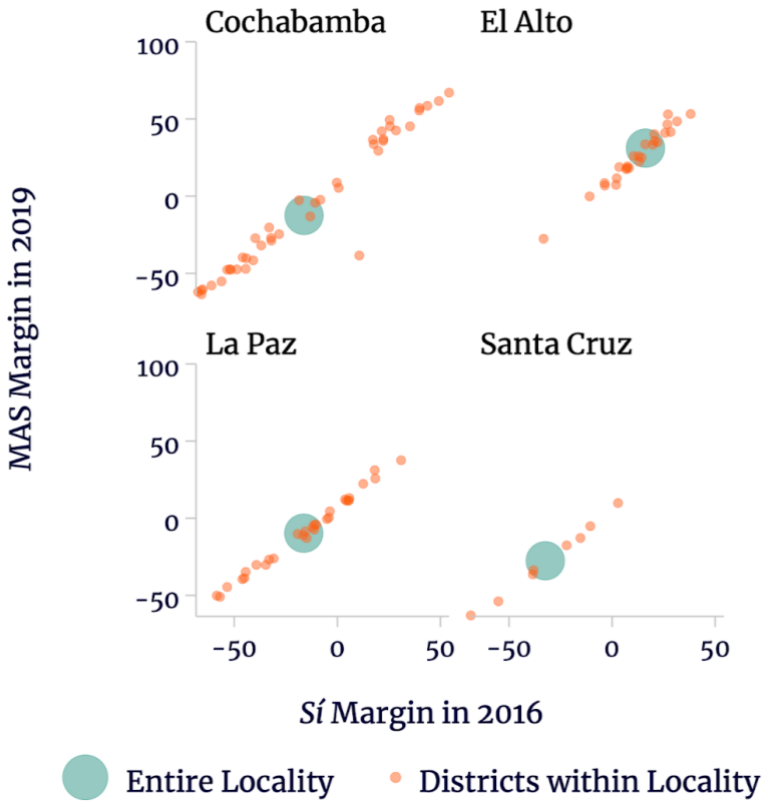
20 OEP (2016b) and OEP (2016c).

21 Provided by email; see Chumacero (2019).

22 Note that from this point forward, we will be referencing the official, final numbers even when referring to polling stations included in the TSE announcement.

Santa Cruz, and La Paz. Still, the results by district varied within each city. Importantly, the more heavily a district favored the referendum, the more the district tended to favor Morales.

Figure 1
Early Polling Stations Suggested a Continuing Pattern of Support Across Elections



Sources and Notes: OEP (2016b), OEP (2016c), OEP (2017), OEP (2019b), OEP (2019c), OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations.

Note that Cochabamba was particularly divided in its support. In both 2016 and 2019, much of the city voted as Santa Cruz had, while other districts voted as had El Alto. The tendency — though imperfect — for patterns of support to be stable across elections suggests that the 2016 results point to some underlying political leanings of the voters.[23]

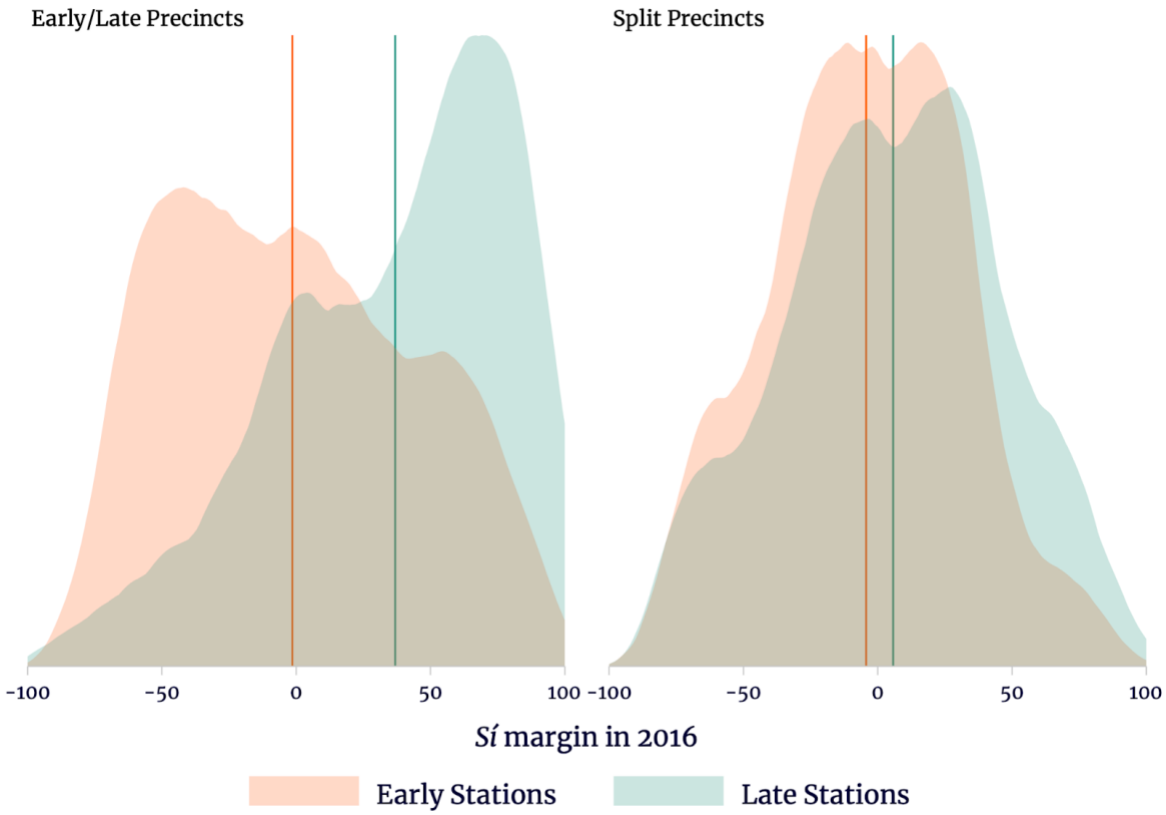
23 The noticeable outlier in Cochabamba is the district of Champa Rancho. Of the 15 polling stations in the district, only the one polling station of the very small precinct (16 eligible voters) of U. E. Alcides Arguedas A. was included in the TSE announcement. U. E. Alcides Arguedas A. was a new precinct in 2019, so the referendum results in its *district* reflect those of the only other precinct in the same district — Unidad Educ. Eduardo Lopez O E. Don Bosco. That precinct represented 3,108 voters across 14 polling stations, favored Morales in 2019, and was counted entirely after the TSE announcement. For whatever reason, the new precinct was almost certainly not representative of the district to which it was added. Thus, the low support for Morales in the early 2019 results of Champa Rancho is not meaningful except that it may cause us to underestimate support for Morales in Don Bosco.

Furthermore, it becomes clear that late polling stations not included in the TSE announcement skewed in favor of the 2016 referendum and thus were highly likely to skew in favor of Morales. For example, in the locality of Cochabamba, 85 percent of polling stations from districts opposing the referendum were included in the TSE announcement, compared to only 60 percent of stations from districts favoring.

In **Figure 2**, we divide polling stations into four groups. In the left panel, we have polling stations in precincts that were either fully early or fully late. In the right panel we have polling stations from split precincts.

Figure 2

The TSE Announcement Drew Heavily from Polling Stations with Unusually Low Support for the 2016 Referendum



Sources and Notes: OEP (2016b), OEP (2016c), OEP (2017), OEP (2019b), OEP (2019c), OEP (2020b), OEP (2020c), Rosnick (2020a), and authors' calculations.

Though the early precincts ran against the referendum by 1.44 percentage points, the late precincts favored the referendum by 36.86 percentage points. In split precincts, those opposed to the referendum were more fully considered in the TSE announcement. Early polling stations in split precincts opposed the referendum by 4.33 percentage points, while late polling stations in split precincts favored the referendum by 5.72. *Nickels before dimes*.

Table 2

The Polling Stations Included in the Announcement Were Not Representative of the Election

	Stations	Eligible Voters	Margin	
		2019		2016†
Early Polling Stations				
From Early Precincts	12,463	2,504,600	9.47	-1.44
From Split Precincts	16,512	3,632,628	6.82	-4.33
Late Polling Stations				
From Split Precincts	4,462	986,459		5.72
From Late Precincts	1,118	191,677		36.86
Total	34,555	7,315,364		

Sources and Notes: OEP (2016b), OEP (2016c), OEP (2017), OEP (2019b), OEP (2019c), OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations.
†2016 margins matched for each 2019 precinct and weighted by valid votes in 2019.

On average, then, the late polling stations favored the referendum by an additional 14 percentage points relative to the early stations. If the swing translated into additional support for Morales, this would imply a first-round victory for Morales. Though this simple analysis doesn't imply a victory was assured at the time of the announcement, it demonstrates how exceedingly premature it was for anyone to call for a second round based on the announced results.[24]

It is not clear that either the electoral authorities or official observers recognized the extent of the problem. TSE President María Eugenia Choque Quispe did allude to a delayed arrival of rural votes but offered no further analysis.[25] Still, according to the European Union's Election Expert Mission (EEM), the announced results were in line with two privately run quick counts (conducted by Víaciencia and Tu Voto Cuenta, a joint initiative of the Universidad Mayor de San Andrés and Fundación Jubileo, respectively) and internal analysis by the OAS presumably carried out by its EOM.[26]

24 It is possible to predict, with confidence, a Morales first-round victory based only on the results announced by the TSE and knowledge of the number of eligible voters at all polling stations. See Appendix F.
25 TSE (2019b). The urban/rural divide in Bolivia is significant, but an inadequate explanation for voting patterns. As we observed, urban El Alto favored Morales. Rural Santa Cruz, Pando, and Tarija favored the referendum, but by much less than rural Chuquisaca and La Paz. Overall, rural precincts not only were underrepresented in the TSE announcement, but even among rural precincts, those included in the announcement were less supportive of the 2016 referendum.
26 EEM (2019), 30.

Absent more detailed data, it is not currently possible to verify that these surveys were consistent with the results presented by the TSE. As reported, Víaciencia showed Morales with a lead of only 4.5 percentage points — already more than 3 percentage points less than in the announced polling stations, and before considering the unrepresentative sample presented by the TSE.[27] The EOM reported the Tu Voto Cuenta results as showing a 5.5 percentage point lead[28]— little better, lagging the then current results by more than 2 percentage points. If these surveys were representative of the entire electorate, they underestimated Morales's final margin of victory, which was sure to grow over the last 16 percent of the count.

The OAS has never made its analysis public.[29]

At first, the TSE did not further update the preliminary results. According to Marcel Guzmán de Rojas, the CEO of NEOTEC (the company hired by the TSE to manage the transmission of results), this was to ensure that the published results didn't change during the press conference.[30] Without controversy, the results on the TREP website were frozen with the 19:40:57 update to maintain consistency between the site and the figures that the TSE announced. Importantly, during this time the process of transmission and the verification of actas continued.[31]

There is no clear record of preliminary counts in Bolivia proceeding to completion. However, the EEM had pushed for more.

> The TSE´s initial plan was to stop publishing TREP results at 8pm on election day with about 80 per cent of results. The reason they gave was that official results would start to be published at that time and they wished to avoid confusion between the two sets of results. The EEM and others expressed concerns about this because if the results were close, 80 per cent is insufficient to give an accurate picture and could be misleading. The TSE appeared to take this on board when on 9 and 10 October they explained to the media

27 Página Siete (2019a).
28 OAS (2019b), 18.
29 As late as November 2020, the OAS was misreporting the results at the time of the announcement, citing a margin of only 7.12 percentage points. OAS (2019b), 17. This figure does not include ballots counted from outside the country. At the time of the announcement, Morales was winning these voters by a margin of 30 percentage points.
30 NEOTEC (2019), 2.
31 It is common to refer to the 19:40 website update as the interruption, or shutdown, of the TREP. Based on the TSE logs, this is inaccurate. Counting only the presidential election, between 19:40 and 20:07 there were 2,022 registration events, 2,120 verification events, and 119 approvals recorded in the TREP. The process continued uninterrupted until 20:07, at which point the TSE interrupted the process of verification. Another 1,284 registration events with the TREP were recorded on election day, but after 20:07. The last TREP entry for October 20 was at 23:58:14, and the first for October 21 at 00:00:37.

and political parties that the results would be given in two tranches. The first at 8pm with around 80 per cent of results, and the second at midnight with around 90 per cent of the results.[32]

There existed a clear tension between risking confusion by reporting the progress of both counts in parallel, and risking confusion with a partial result not representative of the entire electorate. Ideally, the preliminary count would be halted while authorities would explain that the results were not representative and that the election would be close, and would urge patience. However, if the authorities and observers did believe erroneously that the results were in fact representative, the balance of tensions would definitively swing toward a decision to halt the count. Indeed, in announcing that the TREP would be halted, TSE President Choque Quispe said this was to avoid confusion between the two counts.[33]

Complicating matters greatly, at 19:30 — just 20 minutes before the original announcement — Ethical Hacking Consultants (EHC) detected a spike in activity from an internal IP address that was not on their audit list. The TSE had contracted with EHC shortly before the election to monitor election IT. Somewhat predictably, EHC repeatedly butted heads with Guzmán de Rojas' more freewheeling approach to infrastructure management.[34] It is not clear why EHC failed to detect the earlier computer activity. According to EHC, Guzmán de Rojas said he repurposed the existing "BO1" test server to assist the TSE in monitoring the progress of the count.[35] The EOM reported BO1 as having been active all day,[36] and before that active in test runs of the system.[37] In any case, seemingly unaware of the cause and unable to reach Guzmán de Rojas, EHC sent a "maximum alert" to the TSE shortly before 20:00.

According to EHC, the resulting meeting was very tense, with TSE members accusing Guzmán de Rojas of fraud and threatening to call the attorney general. If the authorities had indeed become

32 EEM (2019), 30.
33 Los Tiempos (2019a).
34 Ethical Hacking (2019).
35 EHC did conclude that the server did not contain malicious software, but expressed concern regarding a traffic spike which would peak at approximately 19:45–20:00. Ethical Hacking (2019), 35 and Universidad de Salamanca (2021), 91. Note that throughout the EHC report there is a time zone difference of one hour between events discussed and the corresponding emails and graphics. Panama-based EHC observed daylight savings through November 3 (GMT -5), while Bolivia (GMT -4) did not. The summary of the logs supplied in the OAS's final report showed almost all the activity from BO1 consisted of requests for additional work from transcription operators. EHC reported that Guzmán de Rojas said operators could transcribe as many as two *actas* (polling station records) per minute, meaning that at peak every 30 seconds or so an operator would request new data to transcribe. Likely, the pool of untranscribed polling station data had dwindled, and the now-idle operators were making repeated requests in relatively rapid succession — on average perhaps as fast as once every four seconds.
36 OAS (2019a), 20.
37 NEOTEC (2019), 9.

convinced that the preliminary count had become compromised in the last minutes prior to the announcement, then this would add weight in favor of ending the count entirely. At 20:07, the TSE interrupted the verification process.

The TSE laid out the reasons for the stoppage. According to the EOM:

> At 10:40 p.m., the TSE President informed the OAS technical experts that the audit company had detected a server with an unknown IP address that had connected to the system. Furthermore, she reported that the actions this server had executed were under investigation, as well as the reason why it had failed to be detected.[38]

The EOM argued that the preliminary count must be resumed despite the existence of the BO1 server and its logs, which the OAS would later somewhat contradictorily hold up as evidence of tampering with the count.[39] At noon the next day, the TSE reiterated to the EOM its explanation for the stoppage.[40] In 2020, the Office of the Prosecutor concluded that this alert was the immediate trigger for an unplanned stoppage.[41]

Misinformation regarding the election circulated quickly. With the official count only just underway, Mesa claimed prematurely that he had moved on to a second round.[42] Approaching 1:00 a.m., Marco Antonio Pumari, president of the Comité Cívico Potosinista (Potosí Civic Committee, or Comcipo, an opposition organization) released a series of videos claiming to show fraudulent election materials. Though Pumari retracted those claims just a few hours later[43] after meeting with an observer for the OAS,[44] the videos continued to circulate on social media as recently as May 2022.[45]

---

38 OAS (2019b), 18.

39 Ministerio Público, Fiscalía General Del Estado (2021). It bears repeating that there were no preliminary counts in 2020 or 2021. In the face of concerns regarding computer security, the importance of completing the preliminary count is even more strongly overstated. In any case, the official count proceeded despite acts of violence, including the destruction of three of the nine TEDs.

40 The EOM reported at 3:40 a.m. that the TSE offered an additional explanation: "the TREP interruption was not due to technical issues, rather that the goal that had originally been set had been reached: to verify 80% of the tally sheets." OAS (2019b), 18. Though the OAS presents this as a contradictory explanation, the TSE narrative is consistent. The results on the public TREP website were frozen shortly after the 80 percent mark to preserve consistency between the TSE announcement and the website. The question is if — in light of the EHC alert —the TREP should have been allowed to proceed, or the midnight 90 percent update forgone. That is, the OAS appears to be conflating the pause in reporting with the interruption in the TREP process itself.

41 RTP Bolivia (2020).

42 Mesa Gisbert (2019) and Valverde (2020).

43 El Potosí Video (2019).

44 Agencia Boliviana de Información (2019) and Comité Cívico Potosinista (2019).

45 Silva (2022).

The afternoon following election day, an independent researcher publicized their back-of-the-envelope estimate that Morales would eventually win by nearly 11 percentage points.[46] Either lacking sufficient context, or deliberately misinformed that the result was established, opposition protesters complained that the stoppage would be used to reverse the outcome of the election, essentially delegitimizing the expected result.[47] The prevailing theory was that the electoral authorities shut down public reporting of the count because the results were unexpectedly poor for the incumbent president, and the TSE was buying time to forge substitute documentation containing false results. Reminiscent of Florida in 2000, angry crowds shut down the official count in Potosí, setting fire to the department office.

Under pressure from the EOM and others, the TSE did restart the preliminary count on October 21. When the public website updated at 18:29 with 94.9 percent complete, those polling stations newly included in the count were expected to be much more favorable to Morales, supporting the referendum by 12.4 percentage points. It can hardly have been much of a surprise then that Morales now led Mesa by 10.15 percentage points.

With both the TSE and observers having failed to adequately communicate this possibility, the update added to the chaos. Just as the EEM had feared, the initial announcement had offered a misleading picture of the state of the election. Multiple newspapers in Bolivia had prematurely announced the second round.[48]

The OAS — in its role as observer to the elections — publicly challenged the election results, irresponsibly legitimizing the fears of the opposition. In a press release issued at 21:30 the day after the election, and before all the votes had even been counted, the EOM wrote:

> The OAS Mission expresses its deep concern and surprise at the ***drastic and hard-to-explain change in the trend*** of the preliminary results revealed after the closing of the polls.
>
> [... Today] the TSE presented data with an ***inexplicable*** change in trend that drastically ***modifies the fate of the election*** and generates a loss of confidence in the electoral process.

46 Villanueva (2019a).
47 This same dynamic — accurately predicting results, and arguing that if the predictions held true that this would be a sign of fraud — drove many protests in the aftermath of the 2020 presidential race in the United States.
48 Los Tiempos, for example wrote "el 15 de diciembre habrá una segunda vuelta." Los Tiempos (2019b). A collection of front pages published on the 21st are available at Corzo (2020), 32.

> At an appropriate time, the Mission will issue a report with recommendations ahead of a second round. [***emphasis added***][49]

The EOM expressed confidence that a count of legitimate ballots would result in an outcome other than the apparent first-round victory suggested by the preliminary, and wholly unofficial, results. The EOM was so confident in this declaration that, before the official results had been tabulated, it publicly stated that there would be a second round. No analysis was offered. To date, the EOM has yet to provide any reasoning behind their declarations that day. Its press release solidified the opposition's loss of confidence in the process. As would the OAS, media outlets misreported election results and rarely offered appropriate context for the numbers.[50] By 21:15, opposition protesters had set fire to TED offices in Sucre.[51]

The official count concluded on October 25 with an outcome consistent with the early numbers. Luis Almagro, secretary general of the OAS, declared Evo Morales's reelection to be a fraud — a *coup*, even.[52] But just one day after the original press release, and the second day after the election, the Center for Economic and Policy Research (CEPR) challenged the OAS announcement.[53] Though only a partial analysis was possible, it was already clear that much of

---

49 OAS (2019c) and (2019d).

50 The BBC, for example, misreported:

> Bolivia's Supreme Electoral Tribunal is currently running two separate counts. The quick count, at 95.6% of votes verified, puts Mr Morales ahead of Mr Mesa with a lead of 9.33 percentage points. That is just short of the 10-percentage-point advantage he needs to win outright in the first round. BBC (2019).

In fact, Morales already led by 10.12 percentage points. The 9.33 figure excludes all polling stations located outside the country. This was a frequent error made by observers, as the website reporting the results defaulted to displaying the Bolivia-only numbers (see, e.g., Villanueva [2019b]). On pages 70–71 of their final report, the OAS acknowledges this was a point of confusion, though they likewise erred in this same report. OAS (2019b).

The BBC added:

> **The detailed [official] count shows the two neck and neck**. With 72% of the votes counted, Mr Morales just had a 0.58 percentage point lead over Mr Mesa, making a second round highly likely. BBC (2019). **[emphasis in original]**

In as much as the preliminary count tended to include opposition areas early, this was much more the case in the official count. The latter was very highly geographically ordered, because counting depended on hand delivery of results to department offices. As these offices were in heavily urban capital cities that largely supported the opposition, opposition votes very much arrived early, and generally rural areas that were highly supportive of the incumbent arrived much later. Though the BBC's presentation suggests that Morales was underperforming in the official count, the polling-station-to-polling-station differences across the count were very small overall. Morales would win the official count by 10.56 percentage points.

51 Inocente (2019) and Red Uno de Bolivia (2019).

52 OAS (2019f).

53 CEPR (2019).

the shift came from tallying polling stations coming from split precincts already reported to be more favorable to the incumbent president than those counted earlier. *Nickels before dimes.*

Nevertheless, Morales would not be permitted to finish his current term. The OAS audit team would release a preliminary report very early the morning of November 10,[54] reasserting the allegations of fraud.[55] The audit team would later summarize its report as having "detected willful manipulation of the vote at two levels."[56] But the report does not substantiate the allegation. The report did not point to any *numbers* the teams believes to be altered. There is neither credible evidence that the tally sheets in question were in any way significantly different from "results expressed at the polls,"[57] nor credible evidence that the election's outcome would have changed if a revote at the relevant polling stations had been ordered. The statistical arguments provided in the audit have been proven faulty.[58]

The Fallout

Though unrest in the streets had already resulted in opportunistic acts of violence, recent days already had been marked by serious escalations and threats of violence to both government officials and family members alike,[59] made more serious with the spread of open rebellion by police in several cities starting on November 8. Opposition leader Luis Fernando Camacho, having already demanded Morales's resignation, publicly supported the mutiny, even promising to increase retirement benefits.

An election audit team tasked in such an environment has particular responsibilities — no less so a team directed by international observers such as the OAS. In this case, the OAS had already aggravated tensions on October 21 by openly doubting the official results. The casual publication of preliminary conclusions, in the pre-dawn hours of November 10, proved disastrous.

---

54 OAS (2019e). Luis Almagro tweeted about it at 4:05 AM. Almagro (2019). According to Morales, an assistant to Almagro informed him about the report between 1 AM and 2AM. Mancilla (2021).
55 Columbia University professor Andrew Gelman would characterize the OAS statistical analysis as "a joke." Gelman (2019).
56 OAS (2019a), 11.
57 OAS (2019a), 9.
58 See CEPR (2020a); CEPR (2020b) with special attention to the Annex; and Idrobo, Kronick, and Rodríguez (2022) for a peer-reviewed study.
59 Prominent targets included Vinto's mayor Patricia Arce (AFP [2019]); Esther Morales, Evo's sister (La República [2019]); minister of mining and metallurgy César Navarro (AFP [2019]) and his nephew (EFE [2019]); the governor of Oruro Víctor Hugo Vásquez (Erbol [2022]); and the governor of Chuquisaca, Esteban Urquizo (El Universal [2019]).

Though unable to reach Almagro to voice his concerns, Morales nevertheless kept with his agreement with the OAS and announced new elections and a new TSE.[60] But the opposition was already committed to removal. According to then-vice president Álvaro García Linera, the military chain of command was already broken by that morning.[61]

Victor Borda, president of the Chamber of Deputies (third in the line of succession after the vice president and senate president) announced his resignation. Protesters set fire to Borda's home and kidnapped his brother.[62] Borda declared specifically that he was resigning in order to free his brother, held hostage in Potosí.[63]

Commander-in-Chief of the Armed Forces Williams Kaliman's televised "suggestion" that Morales resign[64] proved to be the end of the line for the current government. Morales and García Linera announced their resignations in the hopes of restoring "societal peace."[65] Morales noted that "dark forces have destroyed democracy."[66]

Stepping down did not preserve their safety. In a press conference two days later, Mexico's foreign minister, Marcelo Ebrard, presented the dramatic story of Morales's escape from Bolivia to find asylum in Mexico.[67] The official Mexican government statement on the decision to grant Morales asylum, too, cites "humanitarian reasons and because of the emergency situation that he faces in Bolivia, where his life and safety are at risk."[68]

The violence not only denied Morales an additional term of office but cut short by months his then-current mandate. Morales — Bolivia's first Indigenous president — had resigned rather than dispatching the military to restore peace. "Thank God," said conservative then-senator Jeanine Áñez as she assumed Morales's role the next day. "The Bible has returned to the Bolivian government."[69]

On November 13, the new minister of the interior, Arturo Murillo, threatened both the former minister of the presidency Juan Ramón Quintana and García Linera's brother, promising a

60 Villamontes (2019).
61 Camacho bragged that his father had made a deal with the military. The named go-between, Luis Fernando López, would become the defense minister under the new regime.
62 El Potosí (2019).
63 ANF (2019).
64 DW (2019).
65 EEM (2019), 36.
66 Collyns (2019).
67 Long & Allen (2019).
68 Gobierno de México (2019).
69 OpenDemocracy (2019).

manhunt.[70] The next day, Áñez issued a decree granting impunity to the armed forces in pacifying the country, leading to the massacres at Sacaba on the 15th and Senkata on the 19th.

The preliminary audit report would continue to resonate. At least three dozen high-level electoral officials were immediately arrested, including the election-day president and vice-president of the TSE.[71] In many cases, the audit report served as the sole basis for criminal charges. Reviewing the cases of four members of the Santa Cruz TED, Human Rights Watch found that prosecutors "have not alleged or provided evidence of any specific actions taken by each of them resulting in electoral fraud."[72] The charges were dismissed in January 2021[73] shortly after the installation of a new, elected, government headed by Morales's minister of economy and public finance Luis Arce.

A Political Standard for Fraud

Shifts in trend happen frequently in elections because ballots are not counted at random; the preliminary count in Bolivia's 2019 elections serves as just one example. The official count in the same election offers an even stronger example. While the TREP relied on electronic transmission of results, creating a bias for early reporting in precincts where cellular access was reliable,[74] the inclusion of any polling station in the official cómputo could not begin until its acta and supporting materials were hand-delivered to the regional TED offices. As the offices resided in capital cities, where opposition support was strong, there was a severe bias for the local, opposition precincts to be counted first. Remote, difficult to reach precincts reported much later. Assertions of fraud must rest on much more than a shift in trend.

The flip side, of course, is that it is impossible to completely disprove the existence of fraud. It may be possible to explain results to a degree that the winner's victory is solidly established. However, it is not possible to prove that every single ballot was validly cast. Nearly every democratic activity of reasonable size is affected by some measure of irregularity, bending of rules, and outright fraud. The number of participants is too large to expect all to behave.

70 Página Siete (2019b).
71 Cuiza (2019).
72 Human Rights Watch (2020).
73 Correo Del Sur (2021).
74 This was a known problem in the preliminary count. To help mitigate, mobile units ferried TREP operators to areas with Internet service. While this certainly helped to secure a more complete preliminary count, results in many rural precincts were nevertheless delayed in transmission.

If the mere existence of fraud delegitimizes an election, then, the entire democratic project is effectively over. Though provable election fraud may be a useful standard for prosecutors, the appropriate standard for delegitimizing a given election must be convincing evidence that fraud changed the outcome.

In the case of Bolivia, it is hard to overstate how unlikely the theories of fraud really are. The official documents are public — as of August 2022, images of the actas are available on the OEP website — and these very clearly point to a first-round victory for Morales. Any theory that the legitimate outcome differed must absolutely depend on the official documents having been forged in large numbers. However, with the TREP operator, the notary, the president of the jury, and any party representatives at each polling station all entitled to copies of the signed acta, potential evidence abounds. The OAS stated that it attempted to compare official actas with copies it found, but to the authors' knowledge did not report any instances of numerical discrepancy. Rather, the OAS found that taking observaciones into account, the cómputo matched on all party results for 34,495 of the 34,555 acta images. Certainly, there was no significant number of specific results challenged by any party.

Thus, contrary to the prevailing theory, politically significant fraud would require documents falsified at each fraudulent polling station in real time and with the consent of any representatives of the damaged parties. It is possible that such fraud would be confined to precincts so overwhelmingly favorable to Morales that neither representatives nor jurors from the opposition observed the counting of the ballots there. However, this would greatly limit the scope of possible fraud.

A Very Simple Projection of Late Results

What could be expected of the late polling stations in 2019? The smaller the geographic area served by an electoral area, the more we expect voters to be similar in terms of socioeconomic status and in their political views. In the United States, we not only expect Alabama voters to differ from those in the District of Columbia, but we also see variation within the District. Voters at Ward 7's Randle Highlands Elementary School voted only 1.7 percent for Donald Trump in 2020, compared to 18 percent at Ward 3's Metropolitan Memorial United Methodist Church.[75]

75 District of Columbia Board of Elections (2020).

We may exploit this to produce very simple projections of Bolivian election results at polling stations reported after the TSE announcement. To begin, we simply assume that voters at the early stations of a split precinct are representative of voters at the entire precinct. This is not exactly true. Within precincts, voters are assigned to stations by surname, which may be associated with socioeconomic status (SES) and voting tendency. If, for instance, stations representing lower SES tended to report late, then the TSE sample would be skewed toward higher SES within-precinct and understating Morales's expected support at the remaining polling stations.[76] Nevertheless, we may proceed. As noted previously,[77] the precinct of Colegio Sebastían Pagador is one example of a split precinct only partially included in the TSE announcement.

Table 3

Final Election Results and Margin Calculations for the Split Precinct *Colegio Sebastían Pagador*

	Morales (A)	Mesa (B)	Difference (C=A-B)	Valid Votes (D)	Margin (E=100(C/D))
Early	1,205	395	810	1,919	42.21
Late	4,941	1,682	3,259	7,861	41.46

Sources: OEP (2020b), OEP (2020c), TSE (2019a), and authors' calculations.

Relative to the early polling station results, Morales underperformed expectations slightly at polling stations reported after the announcement, receiving 41.46 percent of the precinct's late valid votes, compared to 42.21 percent of the early valid votes from polling stations included in the announcement. This difference can be explained by chance, but there may be some reason why Mesa received a slightly elevated share of the valid votes in the late polling stations. While fraud is a possible hypothesis to explain the difference, there are any number of benign hypotheses (including chance) to explain the pattern. To keep these projections as simple as possible, we do not seek to explain the difference.

Applying this same idea to other precincts only partially included in the announcement, we find that in the late stations Morales appeared to overperform slightly. In **Table 4**, we see that across all split precincts, Morales received a margin of 19.54 percentage points, compared to a projection of 18.99. This high level of expected support (18.99 percentage points among 4,462 polling stations, in contrast to 7.88 percentage points at the early polling stations) illustrates how, on balance, these later results came from precincts much more favorable to Morales than those included in the TSE announcement. *Nickels before dimes*.

76 Idrobo, Kronick, and Rodríguez (2022).
77 Rosnick (2020c).

Table 4

The Uncounted Polling Stations in Otherwise Included Precincts Leaned Predictably for Morales

	Stations	Valid Votes	Margin	
			Projected	Actual
Early Polling Stations	28,975	5,155,958	7.88	7.88
Late Polling Stations				
From Split Precincts	4,462	827,823	*18.99*	19.54
From Late Precincts	1,118	153,890	*50.14*	52.12
Total	34,555	6,137,671	*10.46*	10.56

Sources and Notes: OEP (2016b), OEP (2016c), OEP (2017), OEP (2019b), OEP (2019c), OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations.

Unless there are other factors that biased early results in Morales's favor, the difference of 0.55 percentage points in these polling stations limits the potential scope for exploring possible post-announcement fraud there to only 4,500 votes. Even if we fail to seek out alternative explanations, this "fraud" would account for less than 0.1 percentage points of Morales's final margin.

However, another 2.5 percent of voters remained in late precincts not at all included in the TSE announcement. Lacking pre-announcement data there, we cannot apply the same kind of projection as in split precincts. These late precincts were surely even more favorable to Morales than any that had been at least partially included in the TSE announcement. As we saw in Table 2, these precincts favored the referendum by an additional 31 percentage points relative to outstanding polling stations in partially outstanding precincts. If we complete our simple projection by guessing that they also favored Morales by the same additional 31 percentage points, then that would bring our estimated margin in these 1,118 polling stations to 50.14 percentage points — just a little below the actual 52.12.

In total, these very crude projections get us to a final expected margin of 10.46 percentage points, meaning these simple assumptions lead to an expectation that Morales would indeed secure a victory in the first round. It is not clear what kind of reasonable assumptions might suggest an alternative outcome. The geographic pattern of early support for Morales is too clear to justify the OAS decision to add fuel to the fire by declaring that the late results inexplicably modified the outcome of the election.

Escobari and Hoover's Initial Fraud Estimate is Not Credible

It is in this context that Escobari and Hoover in 2019 claimed to scientifically quantify the amount of fraud in the election, writing, "We estimate that the extent of the fraud is ***at least* 2.67%** of the valid votes, sufficient to change the outcome of the election."[78] [***Emphasis added***.] They premised this on the theory that the purpose of interrupting the reporting of preliminary results was to generate fraud. That is, they draw a distinction between "early" polling stations included in the TSE announcement (i.e., before the interruption in public reporting of preliminary results) and the "late" polling stations not considered in the TSE announcement, which they literally refer to as their "fraud treatment group."

In this reckoning, the TREP had been humming along just fine (the EHC alert notwithstanding) and free of fraud. The TREP was stopped to allow bad actors to commit fraud on behalf of Morales, and this explains why Morales's official support at the late polling stations was higher than his support at the earlier polling stations.

In constructing the 2.67 percent figure, Escobari and Hoover assume that the early and late stations are in every way qualitatively identical (including all characteristics of the voters registered) apart from the assumption that the late stations have "fraud" added to them. Thus, any statistical differences between the groups must in their estimation be a consequence of fraud.

They claim that later "fraud" stations favored Morales on net by 29.34 votes relative to the early "control" group. Having assumed away any alternative explanations for this increase, Escobari and Hoover argue this indicates the existence of, at *a minimum*, 29.34 "fraudulent" votes on net favoring the incumbent per late polling station — or at least 163,717 votes.

However, the more the difference between early and late reporting tally sheets can be explained by anything benign (say, a different mix of voter characteristics such as socioeconomic status, or an offhand comment to a particular group — what statisticians refer to as "unobserved heterogeneity"), the less the difference may be attributed to fraud.

78 Escobari and Hoover (2019).

It is important to emphasize that not only do they simply attribute the difference in support to fraud, but they assume it to indicate fraud on behalf of Morales. We call this the "Panic Pete fallacy." Panic Pete is a squeeze toy in the shape of a human head and torso, whose eyes, ears, and nose inflate comically when squeezed. In the case of Panic Pete, we know that the thing that captures attention (eyes, ears, and nose popping out) is not located at the site of the underlying cause (the torso).

Image source: Schylling, Inc.

In election terms, Escobari and Hoover assume that the observed increase in support for Morales is due entirely to fraud favoring Morales in the outstanding polling stations (something unnatural about Panic Pete's head), and not due to votes stolen *from* Morales at the stations included in the TSE announcement (squeezing Panic Pete's torso). Their model does not speak to which of these the alternative explanations might apply, nor does their model rule out any explanation for the difference in support other than fraud. The model only measures the size of the difference.

Escobari and Hoover simply attribute the entire difference to fraud. They then presume to avoid the Panic Pete fallacy by assuming there is no squeeze on the early results.[79]

This is not to say they offer no argument to justify these assumptions. Historically, potential for fraud favoring the incumbent receives heightened scrutiny. To the extent that an incumbent party controls the election process, it may make sense that the incumbent party controls the degree of fraud. However, in a case like Bolivia where the actas are produced at the individual polling stations, it is wrong to presume that the incumbent party is the more likely generator of fraud in areas dominated by the opposition.

That there can be no alternative explanation but fraud for the difference is an obviously implausible assumption to make. There is simply no question of the late polling stations having exhibited much greater support for the referendum than the early polling stations. Ultimately, it is unclear what kind of assumptions could credibly explain a 2.67 percentage point fraud.

79 This applies equally to our above projections of the final count based on the pre-interruption results. Though we say that Morales overperformed (slightly) at polling stations not included in the announcement, this does not mean that we believe the early results were natural and the late results artificially inflated. It could be that the early results were artificially depressed. The EHC "maximum alert" was triggered by events that took place before the TSE announcement, so there could have been fraud there — squeezing the torso to make Pete's eyes bug out. Or, there could be some other benign factor that such a simple model for projecting votes fails to capture.

As we have already observed, the outstanding stations more heavily favored the 2016 referendum. Combined with the early results, this explains almost exactly the difference in early and late results, and it is demonstrably implausible that fraud explains much of the 164,000 votes. In any case, it is simple to show that 2.67 percent is not a credible estimate of fraud in the election.

To start, we note that among the 4,462 late polling stations in split precincts, Morales outpolled Mesa by 161,791 votes. If the alleged fraud is evenly distributed across late polling stations, then this margin of victory is almost entirely wiped out as only 30,876 votes presumed legitimate remain there.[80] As a result, the resulting "fraud-free" margin on these polling stations would be only 3.73 percentage points — about 15 percentage points below what we might reasonably expect based on the early results in these precincts. For the 2.67 percent figure to be credible, the alleged fraud must lie overwhelmingly in the 1,118 polling stations that have no early results — those in the late precincts.

In **Table 5**, we artificially reduce Morales's margin of victory by a total of 163,717 votes. As the early results are assumed to be free of fraud in Escobari and Hoover's estimates, we do not adjust these. We adjust to the late polling stations from split precincts by 10,303 votes[81] — more than twice the adjustment of Table 4, or about 1.24 percentage points. This requires finding another net 154,414 votes for Mesa in the few remaining precincts that were entirely excluded from the TSE announcement. As there were only 153,890 votes in these precincts, the official results would have to be very far off indeed.

Table 5

Escobari and Hoover Require Massive Fraud in Late Precincts

	Valid Votes (A)	Net Support			
		2016 (B)	2019 (C)	2019 Hypothetical (D)	Difference (E=D-B)
Early Polling Stations	5,155,958	-3.17	7.88	7.88	11.03
Late Polling Stations					
From Split Precincts	827,823	5.77	19.54	18.30	12.53
From Late Precincts	153,890	36.86	52.12	-47.57	-84.44
Total	6,137,671	-0.97	10.56	7.90	8.85

Sources and Notes: OEP (2016b), OEP (2016c), OEP (2017), OEP (2019b), OEP (2019c), OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations.

80 That is, 161,791 - 29.34 x 4,462 = 30,876.
81 This is 4,462 x 2.309 ≈ 10,303, where 2.309 is Escobari and Hoover's separate estimate of the average excess of Morales votes in these polling stations. For the sake of argument, we set aside the credibility of this estimate.

For the "fraud-free" election result to be at least 2.67 percentage points lower than the official 10.56, Mesa must have outpolled Morales in these precincts by more than 47 percentage points. Instead, the official results show the reverse: Morales outpolled Mesa by 52 percentage points. Indeed, while the early results favored Morales by 11 percentage points relative to the 2016 referendum, Escobari and Hoover imply that in an honest accounting, Morales underperformed the 2016 referendum by 84 percentage points in the late precincts.

Escobari and Hoover's allegation of fraud is therefore not a matter of adding a few votes here and there, but an obvious and massive reversal of results in precincts where they claim voters must have overwhelmingly supported Mesa.

Escobari and Hoover offer no explanation for how these precincts, supposedly favoring Mesa by large numbers and with counts surely witnessed by opposition supporters, could have generated results overwhelmingly in favor of Morales with neither opposition representatives retaining copies of the results, nor Mesa's team officially challenging the results.[82] Escobari and Hoover's interpretation of their 2.67 percent is simply not credible.

The obvious alternative explanation for the 2.67 percentage point difference is that the late polling stations represented very different people than the early polling stations, and their model simply measures bias in selecting polling stations for the initial announcement of results. As far as we know, there is no dispute over the fact that the late polling stations were historically and predictably favorable to Morales. We thus utterly reject any claim of late-station fraud on this scale.

Absent further evidence respecting the late precincts, Escobari and Hoover could have presented 0.17 percentage points as a reasonable estimate of unexpected votes for Morales, recognizing

82 Escobari and Hoover later cite Chumacero, saying he reported 99 percent of accessed images of the tally sheets for polling stations never publicized in the preliminary count, and that they "contained discrepancies between the pictures and what was reported in the official results." Escobari and Hoover (2020), 19. However, there is good reason to believe these matter much less than it might appear. For example, Chumacero writes "*en 12% de ellas hay más votos que personas inscritas*" ("and 12% recorded more valid votes than the number of people registered to vote," as translated in Escobari and Hoover [2020], 19). This language is unclear. There are precisely zero polling stations for which the official count recorded more valid votes than eligible voters. Chumacero likely meant *emitidos* — valid votes, plus blank and null votes. However, there are still zero such stations in the official results even when blanks and nulls are included. Most likely, Chumacero is counting clerical errors on the part of the jurors. Most, if not all, of these polling stations were not included in the preliminary count because there were unresolved problems apparent in the images, including clear evidence of clerical errors, whether observed or missed at the corresponding polling station. The votes at such polling stations were reconsidered in the official count. In accordance with electoral law, the TED reviews all the materials in the electoral package to resolve such errors before resorting to nullification of some or all results at the polling station. There is nothing to indicate significant discrepancies between the tally sheets considered in the final count and official results.

that the model may not control fully for all benign factors affecting the difference in votes. This would of course fall far short of reversing the outcome of the election. Instead, they dismissed their own result, arguing that the model fails to pick up fraud in the late precincts and that there must be at least 2.67 percentage points of fraud. As we saw in Table 5, it is not credible that there is sufficient fraud in the late precincts to make up the difference.

Escobari and Hoover's Updated Paper

In 2020, just days before a new presidential election was held in Bolivia — this time with the former president in exile — Escobari and Hoover released a greatly expanded version of their paper. The update only slightly reduces the scope of the fraud allegation — now 2.50 percent of valid votes, down from 2.67. This 2.50 percentage points of fraud in late polling stations is hardly more credible than 2.67. However, Escobari and Hoover now claim fraud throughout the count, with three-quarters taking place in the early polling stations included in the TSE announcement. That is, they argue that the announcement itself overstated Morales's actual support at the time by about 2.3 percentage points.

Recall that Escobari and Hoover's original estimate was premised on the idea that the early votes served as a "fraud-free" control by which they might contrast the late polling stations contaminated by fraud. By allowing the possibility that fraud took place throughout the election, they have no real theory of fraud running one way or another. They abandon their justification for the TSE announcement as one dividing "control" (fraud-free) from "treatment" (fraud-contaminated) polling stations. This fully reinstates the Panic Pete fallacy. If fraud is possible anywhere along the line, Escobari and Hoover make no effort to distinguish fraud in favor of Morales from fraud against him.

Even granting, generously, that any possible fraud is one-sided, Escobari and Hoover face one fundamental problem common to all their analyses. Just as coin size relates to bias in selection order as well as value, geography — down to precinct locations — is a confounding variable that affects both support for the incumbent (Morales) and the time when votes are counted. That is, for various reasons, much of Morales's support was counted relatively late. Failing to fully account for this leads to a bias in their results and thereby generates the appearance of statistical "evidence" of fraud, as we will document in subsequent sections.

Further, Escobari and Hoover have repeatedly denied researchers access to their data and methodology. Their principal contribution in their update could have been to match polling

stations in the 2019 election with polling stations in the 2016 constitutional referendum. It is unclear that it is possible to match tally sheets reliably across elections, given changes in precinct boundaries, though Escobari and Hoover claim to have done exactly this. However, their summary results cast serious doubt that they matched the tally sheets with sufficient care. They claim to have identifiers that allow matching, but it is highly unlikely that these identifiers match properly across elections. In particular, the 2019 election had 34,555 tally sheets — 33,048 from within the country, and 1,507 representing voters abroad. By contrast, data from the Órgano Electoral Plurinacional de Bolivia (OEP) shows that the 2016 election had only 29,224 tally sheets within the country, and 1,143 abroad.[83]

If this mismatch poses an insufficient obstacle, consider that the total number of tally sheets across the two elections is 64,922. Yet Escobari and Hoover report across Tables 3 through 6 no fewer than 65,811 observations. Not only do the tally sheets not match across years, but there are more observations than total tally sheets. We have been unable to reconcile these numbers, and repeatedly Escobari and Hoover have refused to answer questions.[84] Thus, we will not attempt to replicate their findings fully.

Instead, the balance of this paper consists of two sections. First, we will present some stylized facts about the election results with an eye to constructing synthetic data sets over which we will have complete control of the degree of "fraud." Then, we will apply Escobari and Hoover's models to these synthetic data sets, illustrating how they fail to capture fraud, and instead measure bias in the ordering of polling stations — *nickels before dimes*. In Appendix A, we will examine their regression discontinuity result.

Some Stylized Facts

In the successful estimation of the late votes in split precincts, we observed that voters attending the same precinct tended to vote similarly in 2019. In the prediction of the vote at late precincts, we relied on stability of vote patterns across elections. It is well established that certain areas

83 OEP (2016a).

84 While we recognize that they have made public only a working paper, and not a study for publication, it is not simply some ivory-tower exercise. Working papers allow researchers to share ideas. Much academic work may depend on holding data private until formally published — lest a better-connected researcher scoop a working paper's authors. In this case, Escobari and Hoover weighed in (again) on an explosive public debate. Regardless of the merits of publishing working papers without data, a higher standard of transparency is certainly required here. Instead, on the very day it was updated — just prior to the 2020 election — Escobari and Hoover found their unverifiable conclusions reported in the press as having established fraud; see e.g. Página Siete, (2020a). Just as with the OAS, public claims demand public data.

tended to vote one way or another. As noted earlier, the capital cities voted heavily for Carlos Mesa. El Alto — the second-largest city in Bolivia — voted strongly for Morales. Rural Bolivia also voted heavily in favor of the 2016 referendum and for Morales. The city of Cochabamba was particularly divided, geographically, with districts showing widely varying levels of support in 2016, but each district showing little change from 2016 to 2019.

We summarize this notion, that voters are more similar the smaller the geographic region, by saying that support for the referendum and incumbent are connected to "geography." Geographic differences include broad SES, literacy, rurality, and any number of expressly political factors such as whether a school was built in the area, or a mine explored (or left unexplored). A candidate's success or failure in pronouncing the name of a community or a person in the community may impact local results. Any of these factors can help explain support for a given candidate in 2019, or a change in strength of support across elections.

Importantly, geography may also be associated with the "arrival" of an acta to the TREP. That is, there is something connected to the geography that also associates with the order in which results were transmitted. As discussed above, for example, there was a transmission bias in the TREP due in part to variation in cellular phone coverage, delaying results from remote parts of the country. It is also possible that in some more rural areas, where Spanish literacy is lower, that jurors were slower to complete actas, say, due to requests for assistance. Community democracy, practiced in many rural areas, introduces a very highly geographic component to the elections.[85] In other words, factors that may influence voter preferences may also influence "arrival" order.[86]

Through most of the TREP, *at least* 1,000 actas — almost 3 percent — were transmitted every five minutes.[87] Thus, the order in which actas arrived was very sensitive to almost any kind of noticeable delay: a five-minute delay could at peak move an acta 1,700 places later in the order. We see this in **Figure 3**. The "ARRIVAL" of an acta here is defined in terms of the share of actas transmitted: the first acta to arrive is assigned an "ARRIVAL" value of zero, while the last is assigned a value of 1.[88]

---

85 As stated by the OAS (2019b), 213:

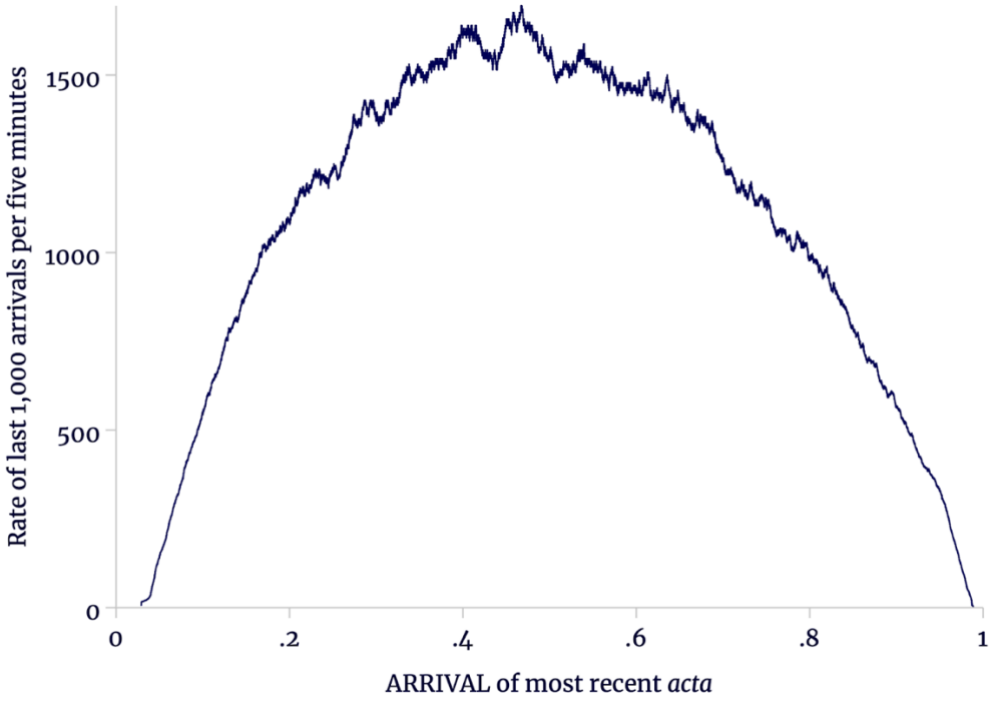
> The 2009 Constitution establishes a plurinational State that recognizes the thousand-year roots of the indigenous peoples and includes community democracy as one of their forms of government, which is exercised with respect for the norms and procedures unique to the original rural-dwelling (*campesino*) nations and peoples.

86 OAS (2019b), 26.

87 Again, the image and transcript were transmitted separately. The data from the TSE reports the time that the transcript arrived electronically at the TREP system.

88 Specifically, we order actas by Ultimate Transmission Date (in seconds). We can see that 394 actas were not transmitted, and were placed at the end.

Figure 3
Rate at Which Actas "Arrived" in the TREP



Sources and Notes: Rosnick (2020a), TSE (2019a), and authors' calculations.

In terms of counting coins, we had two types. The larger coins get picked first, and have lower value. Over time, then, the average value of the coins selected trends upward, but it is not that the smaller size of the dime drives the higher value. It is also not that the dimes' greater value led them to get picked later. Rather, the greater value and smaller size are both properties of dimes, so the bias in selecting large coins first — which happen to be nickels — is also a bias in favor of selecting lower-valued coins first.

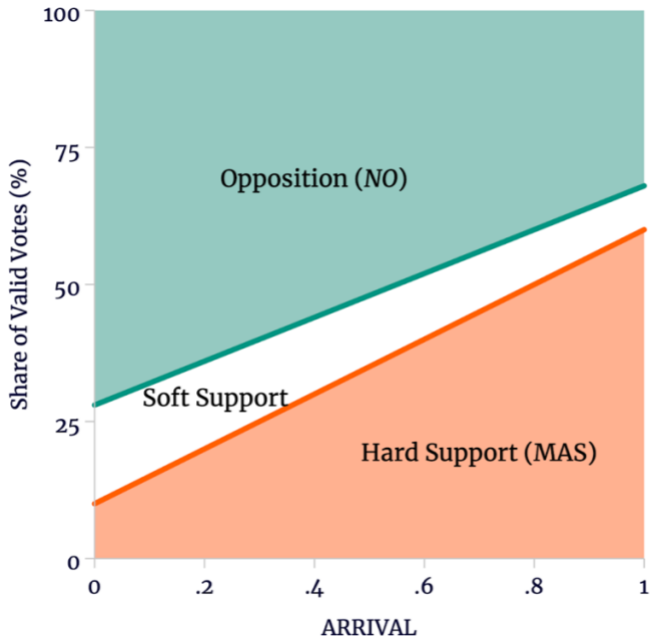
Likewise, in Bolivia, delays in transmission are more likely in areas favoring both the referendum and Morales — creating an early bias in favor of the opposition. We know that polling stations favoring the referendum in 2016 tended to be counted late; the early results at the time underestimated support for the referendum in polling stations that reported later.

The Effect of Geographic Bias Need Not Be Stable Across Elections

To the extent that "geography" can explain this bias, we may expect geographically induced bias to exist in 2019. It may be possible that some of the bias in arrival order is reduced from 2016 to 2019 — perhaps due to increased cellular coverage. But as not all factors are well-known, this is not well-quantified. The bias could be worse in 2019.

Importantly, the impact of the bias on the overall difference in support between early and late polling stations may be larger even if the bias is smaller. This is because the 2019 results may simply be more sensitive to geography than those in 2016. Such a difference in sensitivity may have a completely benign *nickels before dimes* explanation.

Figure 4
Hypothetical Shift in Support Between Elections



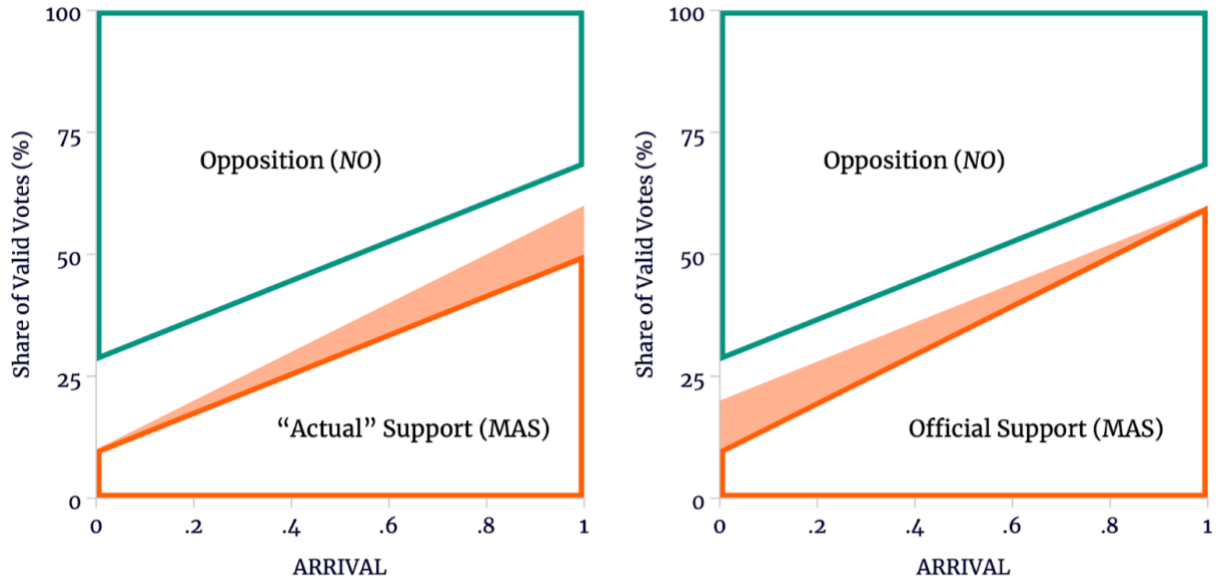
Sources and Notes: Authors' calculations.

Then in effect, the opposition will have peeled away 18 percentage points from Morales in the earliest-arriving polling stations, but only 8 percentage points in the latest-arriving polling stations. The narrowing difference between 2016 and 2019 is not caused by the order of arrival itself, but that the support for Morales is wider and harder among the late arrivals, leaving a smaller pool of voters upon which the opposition can feasibly draw.

In this example, support is more sensitive to arrival in 2019 than in 2016 because soft support is sensitive to geography. This is true for entirely benign reasons, and we see a difference in trends within the historical data. Escobari and Hoover present similar results.[89] As we will discuss later, they simply assume that this narrowing of soft support is an indication of fraud — that despite the above illustration, there can be no benign explanation. Worse, they assume that it indicates adding net MAS votes among late arrivals, rather than the equally plausible theft of MAS votes among early arrivals.

Figure 5

Even If the Slopes Should Be Identical, the Direction of "Fraud" is Ambiguous



Sources and Notes: Authors' calculations.

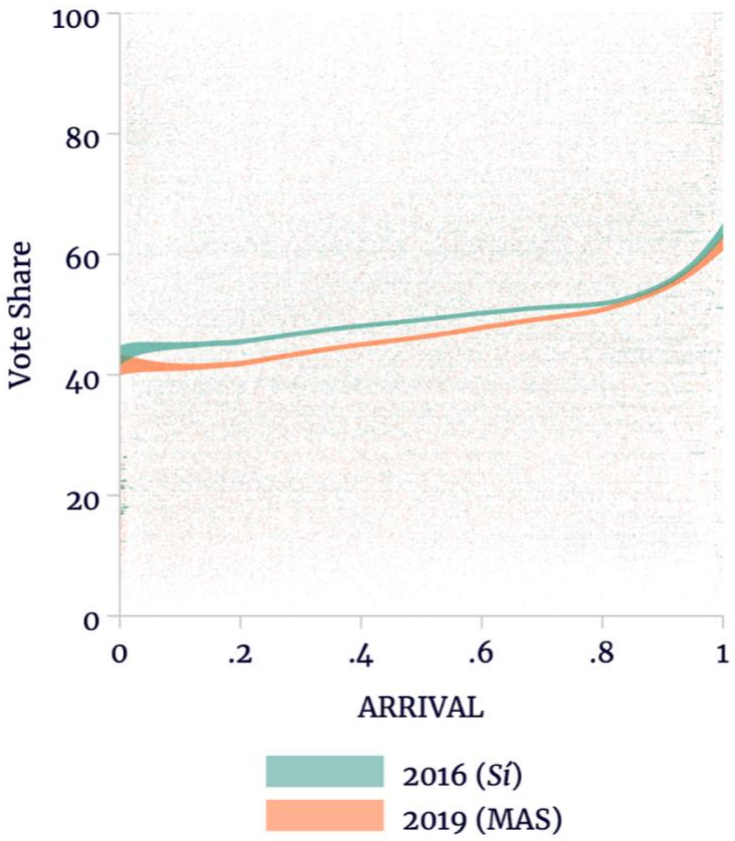
On the left, we have Escobari and Hoover's preferred interpretation: the shaded orange area reflects official support for MAS that they presume to be fraudulent. On the right, we have an equally valid interpretation: the shaded area reflects actual support for MAS fraudulently taken by the opposition and therefore merely appears soft in the official results.

In either interpretation, the actual soft support (voters in favor of the referendum, but not of Morales) is a constant share of the valid vote regardless of arrival. The difference is merely one of size. There is no reason to think that the soft support should follow either pattern of **Figure 4**, and so no reason to believe the different slopes in 2016 and 2019 indicate fraud.

89 Escobari and Hoover (2020), Figure 4. See Appendix B.

For the moment, we simply confirm that the data suggests a narrowing of support. In **Figure 6**, we present similar results to those seen in Escobari and Hoover.[90] Our figure shows the MAS/*Sí* share of the valid vote for every polling station,[91] ordered by arrival.

Figure 6
Actual Trends in MAS/Sí Support Over Arrival



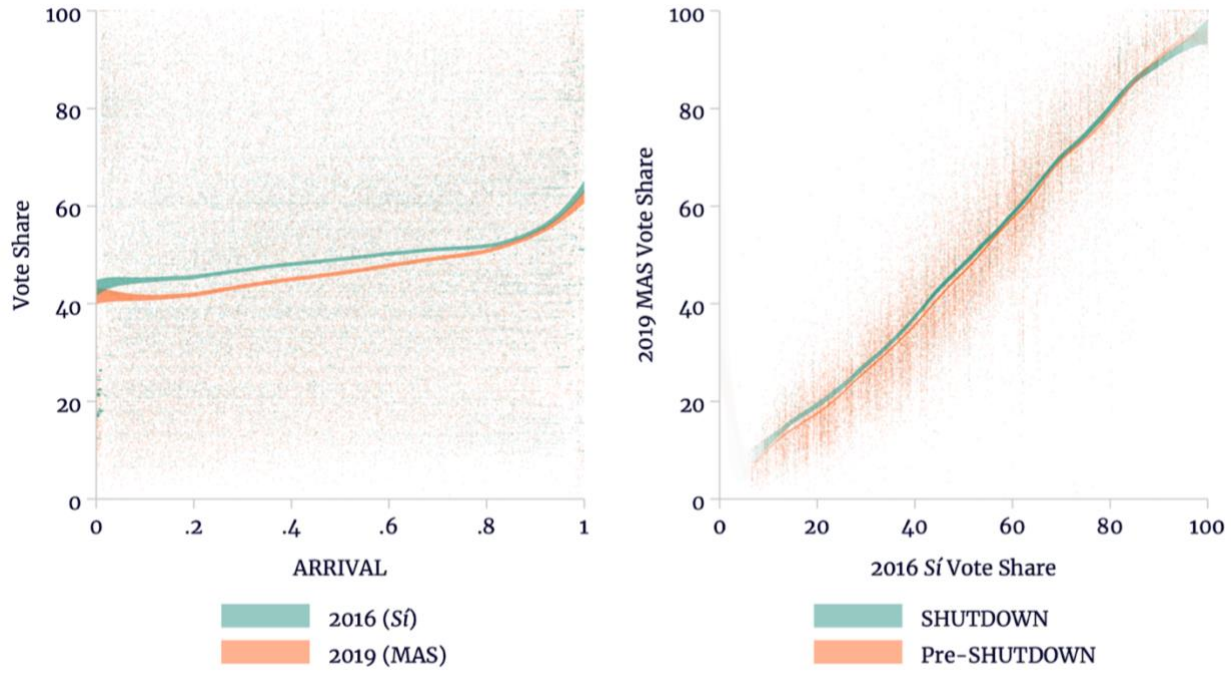
Sources and Notes: OEP (2016b), OEP (2016c), OEP (2017), OEP (2019b), OEP (2019c), OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations.

Two features of the data stand out. First, though the MAS share is generally lower than the *Sí* share, both trend upward, picking up among later arrivals. Second, the difference between the two narrows over arrival, nearly converging by the end. For the moment, we simply observe that this is the case. Later, we will examine the effect this has on the arguments presented by Escobari and Hoover.

90 Escobari and Hoover (2020), Figure 4 (left panel).
91 As discussed above, 2016 shares are generally at the precinct-level, but in some cases the geographic match is wider than the precinct, say district, locality, or municipality, as required.

In **Figure 7**, we place Figure 6 alongside an alternative way of presenting the data. Rather than arranging polling stations in order of arrival, we arrange them by the matched support for the referendum. Rather than dividing by election, we separate those late polling stations not included in the TSE announcement. In keeping with Escobari and Hoover, we label these "SHUTDOWN" polling stations.

Figure 7
Two Ways of Looking at the 2019 Results



Sources and Notes: OEP (2016b), OEP (2016c), OEP (2017), OEP (2019b), OEP (2019c), OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations.

Looking at the panel on the left, Escobari and Hoover explain the gap in early arrivals as "consistent with Morales's known loss of popularity between 2016 and 2019."[92] The implication is that the later arrivals — lacking this gap — must include fraudulent votes. For Escobari and Hoover, the post-interruption polling stations, with their higher MAS shares, are suspect.

However, the graph on the right suggests that there is almost no difference on average between pre- and SHUTDOWN results when arranged by support for the referendum in 2016. Polling stations with the same level of support for the referendum showed nearly the same level of

92 Escobari and Hoover (2020), 16.

support for Morales regardless of whether the polling station reported before or after the interruption: the difference seen on the left reflects a bias of *nickels before dimes*.

This does not mean there was no fraud in the late polling stations. The election was too closely decided to make any determination based on these graphs. However, it appears that if there is indeed any unexpected support for Morales after the interruption, it lies in precincts where the referendum was relatively opposed — not strongly supported. If anything, where the referendum was most heavily favored, Morales underperformed in the late reporting polling stations.

So, we have observed that geography can play a role in the ordering of the count via small delays in cellular coverage or slowness in counting or filling the acta. We have observed that the effect need not be consistent across elections. There are other observations that warrant notice regarding the timing of the reporting of results.

First, actas containing observaciones may have come in late, as well as have been processed late, due to delays within the TREP itself. The time needed to make such notes and corrections may delay transmission.

Second, after jurors complete their acta, they still must wait on the availability of a TREP operator to transmit the results. As there was only one operator for every 10 polling stations, this limited how quickly results could be sent to the SERECI.

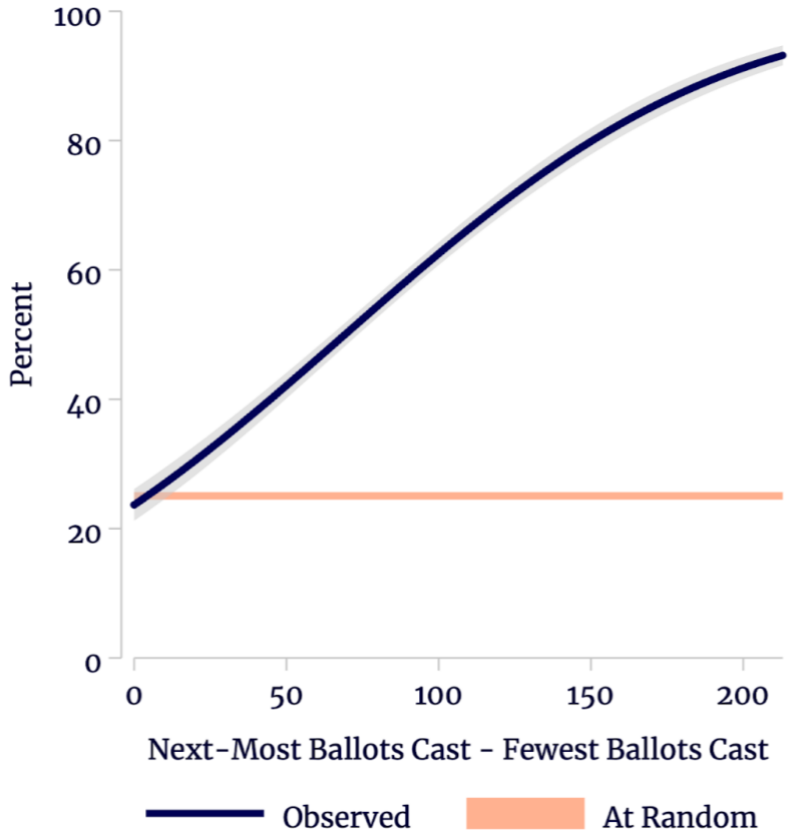
Third, there was a disproportionate tendency for tally sheets with relatively few numbers of votes to report very early. This makes sense, as with fewer votes to count, less work is required, and the count at such ballot boxes may be completed quickly.

Of course, if the polling station with the fewest ballots has many fewer ballots than the station in the precinct with the second-least number of ballots, then it will be very likely to report first. If the number of ballots in the two stations are very close, then it should be nearly random as to which polling station in the precinct reports first. This is what we observe in **Figure 8** below. As the gap in ballots cast grows, the more likely it is that the precinct's smallest station reports first. As the difference becomes small, the probability is close to 25 percent.[93]

93 As the number of polling stations in each precinct has a harmonic mean of 3.9–4.1, we expect that on average 1 in 4 of the "smallest" polling stations would report first if at random.

Figure 8

Probability That the Polling Station with Fewest Ballots Reports First



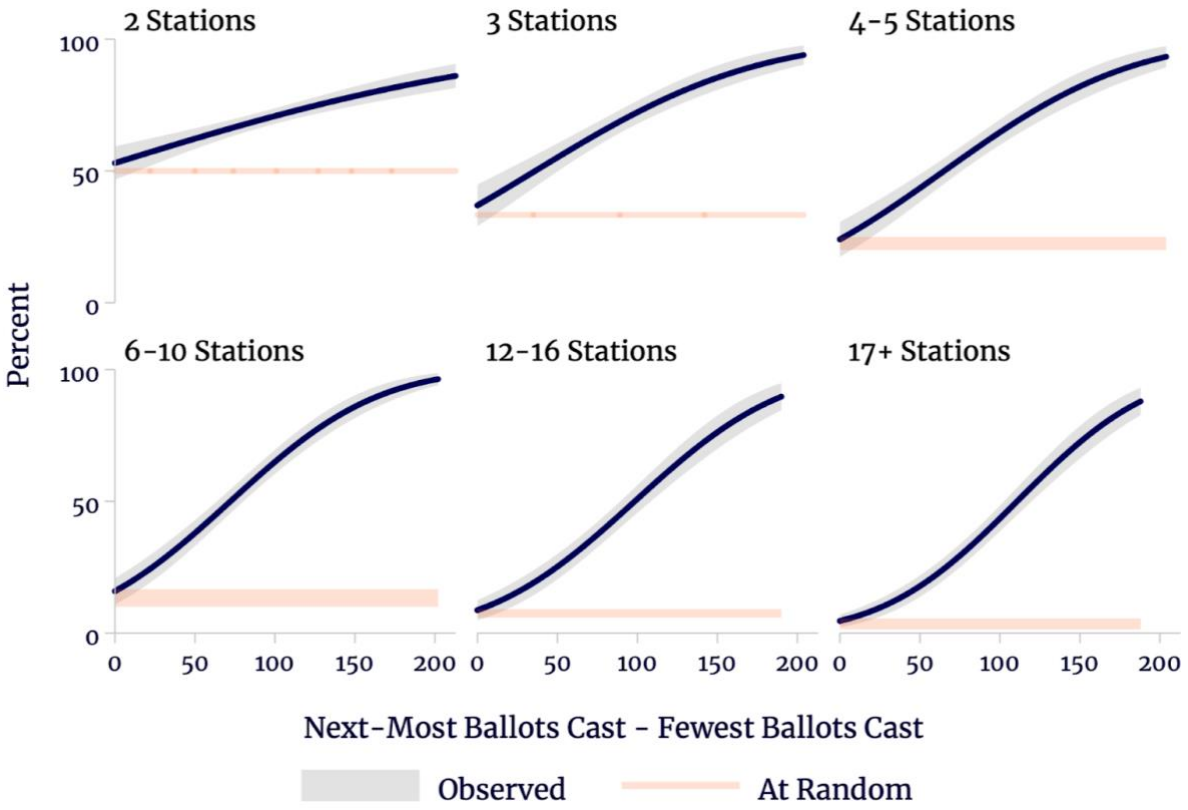
Sources and Notes: Rosnick (2020a), TSE (2019a), and authors' calculations.

This relationship holds if we separate urban from rural, or observe by number of polling stations, as in **Figure 9** below.[94]

94 Note that in their original paper, where Escobari and Hoover propose a minimum amount of fraud detected in the election, the result comes from an analysis of the differences in candidate vote totals in each polling station, rather than in shares. Given that smaller stations report earlier than larger ones, this model is inappropriate. To illustrate: suppose that every polling station is won by one candidate or the other by two votes to one. Suppose Morales wins 60 percent of them, and Mesa only 40 percent, regardless of when the polling stations report. Now suppose that some stations have only 75 votes, while the others have 225. If the smaller stations report first, then the average difference in votes per polling station rises by 10 votes, though size (and therefore arrival) are the only features to distinguish them. Labeling this "fraud" is plainly wrong.

Figure 9

Probability That the Polling Station with Fewest Ballots Reports First — By Number of Stations



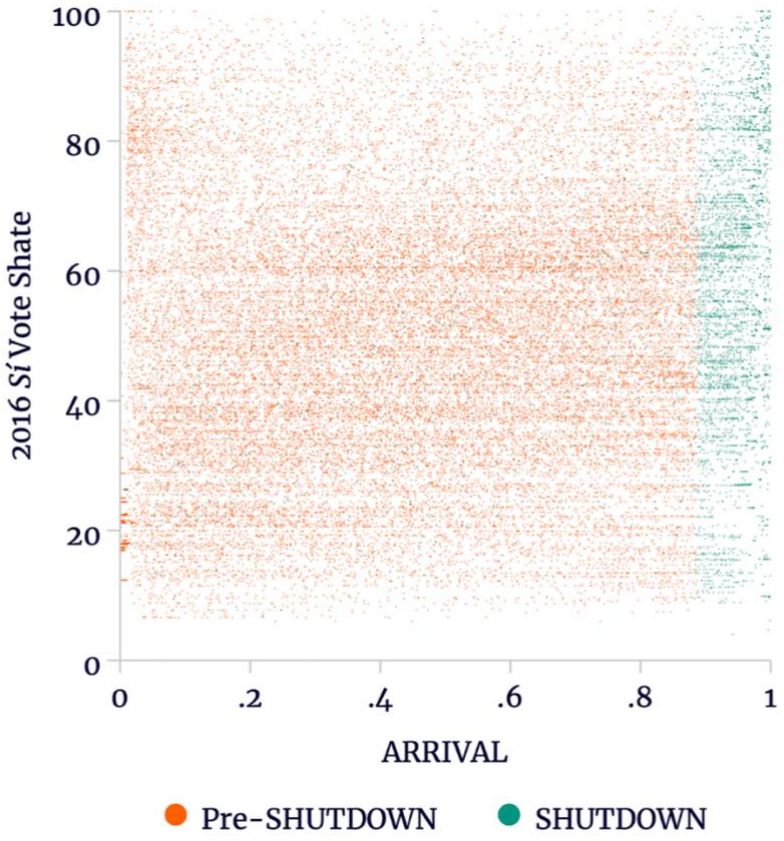
Sources and Notes: Rosnick (2020a), TSE (2019a), and authors' calculations.

Two-hundred of the 5,132 tally sheets with the fewest ballots to count were not timestamped as transmitted in the TREP, and 194 of these 200 were rural. Of these 194 precincts, 156 had only one tally sheet. Of the remaining 38 precincts, 30 had only two tally sheets. Twenty-nine of these 30 had neither tally sheet timestamped as transmitted. Of the eight precincts remaining, these had between three and six tally sheets — none of which were transmitted. In other words, these were small, rural precincts that largely failed entirely to report in the TREP.

Finally, we observe that although some tally sheets arrived early but were not included in the TSE announcement due to transcription issues, most polling stations excluded from the announcement were simply among the last to arrive.

Figure 10

Arrival Order of Polling Stations



Sources and Notes: OEP (2016b), OEP (2016c), OEP (2017), OEP (2019b), OEP (2019c), OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations.

Constructing Synthetic Data

Again, Escobari and Hoover's summary tables raise questions regarding the replicability of their results, and their failure to respond to questions make such an exercise extremely difficult. To illustrate some of the underlying problems with their analysis, we turn to synthetic data. The synthetic results will meet some of the above stylized facts, but as these are meant to be illustrative, the underlying trends will be clearer and simpler.

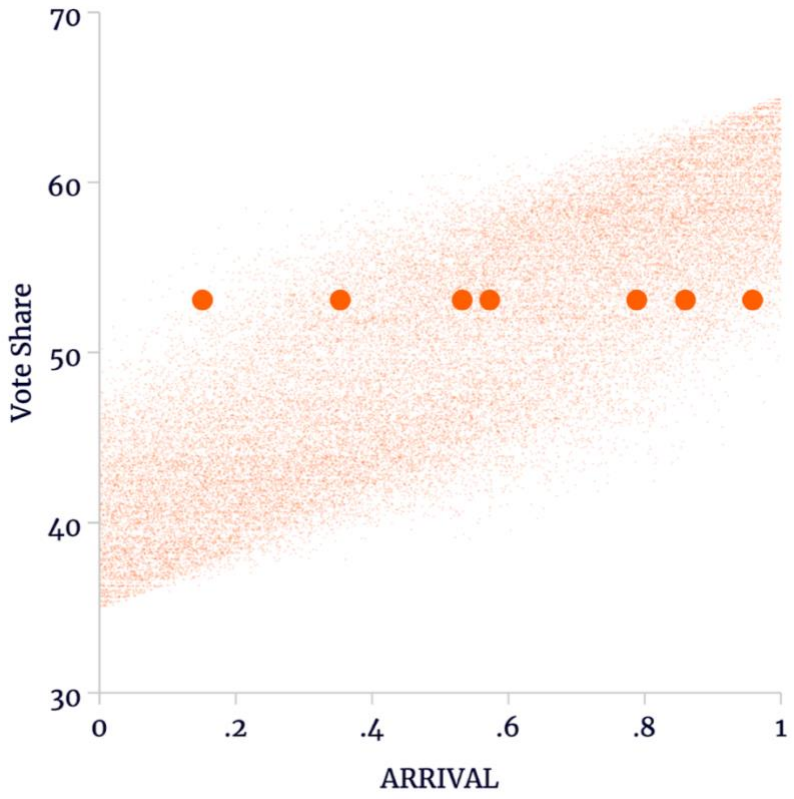
Importantly, we control the amount of "fraud." In doing so, we see that applying Escobari and Hoover's methods to perfectly clean election data will consistently fail by erroneously inferring fraud where none exists.

We begin the process of generating synthetic election results with a set of precinct-level values between 0 and 1 and with a mean of 0.5, which we call "geography." These values represent geography in an abstract sense so that precincts with values close to one another have voters more alike than precincts with values more distant from one another. Polling stations within each precinct share the same "geography." Therefore, within any precinct, voters are effectively identical irrespective of the polling station that they attend.

In this data, we construct geographies to vary with the arrival order of the synthetic polling station election results. In this way, we see something about the voters (represented by geographic variable) that affects the timing of the vote count.

Figure 11

Example Synthetic Data with No Causal Relationship Between Transmission Order and Vote



Sources and Notes: Authors' calculations.

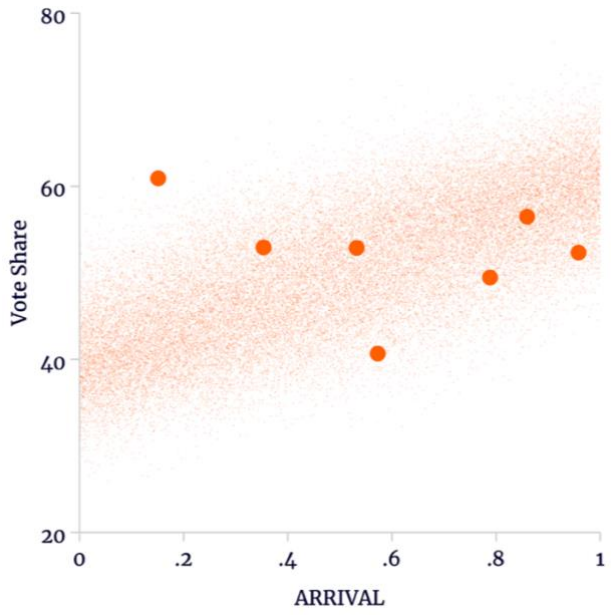
We may then assume some additional relationship between geography values — again representing voters — and precinct-level support for a candidate or a position. Thus, in our data geography relates to both vote share and (in part) to arrival order, yet neither does vote

preference cause arrival order, nor vice versa.[95] Nevertheless, through geographic effects, vote preference is associated with arrival order. Finally, we assume that voting preferences for a particular candidate correspond to vote share for that candidate — that is, there is no fraud. In **Figure 11**, we see expected *precinct*-average support for a candidate given assigned geography. Every polling station within each precinct is marked at the station's individual arrival.

In Figure 11, one precinct is highlighted to demonstrate that the trend is entirely due to variations across precincts. Polling stations within any given precinct are identical in their expected vote share. Late polling stations within a precinct are not associated with higher vote shares.

However, if support for a candidate is 40 percent throughout a precinct, that does not imply that the candidate will receive exactly 40 percent of valid votes at every polling station in the precinct. Even if voters were randomly assigned to polling stations, pure chance will vary the level of support observed at each station so that results might look something like **Figure 12.**

Figure 12
Results May Vary Widely Among Polling Stations at Individual Precincts



Sources and Notes: Authors' calculations.

95 This is a hypothetical, so we assume that causality runs from geography to voting. This is a convenience. Reverse causality — that voting preferences lead people to live in certain areas — is also possible. There may be reasons not overtly political for an area to attract residents of a particular political view. Over time, the area becomes associated with people of a specific orientation. It is unimportant in this analysis if ultimately voting preferences cause geography. Indeed, we might even assume that people of a particular persuasion move to rural areas with poor cellular coverage, and so report late. We could even assume some people intentionally move to an area that reports earlier or later for the purpose of seeing their votes reported relatively early or late. Ultimately, all that is important is that arrival order is associated with actual political preferences, for whatever reason.

Here, the vote shares within a precinct are no longer identical. Within a given precinct, there may be a trend upward or downward, but this is entirely due to chance and not to fraud. Across many precincts, the average noise-driven trend approaches zero.

Importantly, the results in Figure 12 are observable, in the sense that we would know when a tally sheet (representing a polling station) arrives, and what the actual votes recorded on that sheet are. We would not be able to observe directly that the vote shares do in fact correspond to voter preferences, but it is possible to analyze the election equipped with this data alone, with no prior knowledge of the (unobservable) geographic variable that generated the data.

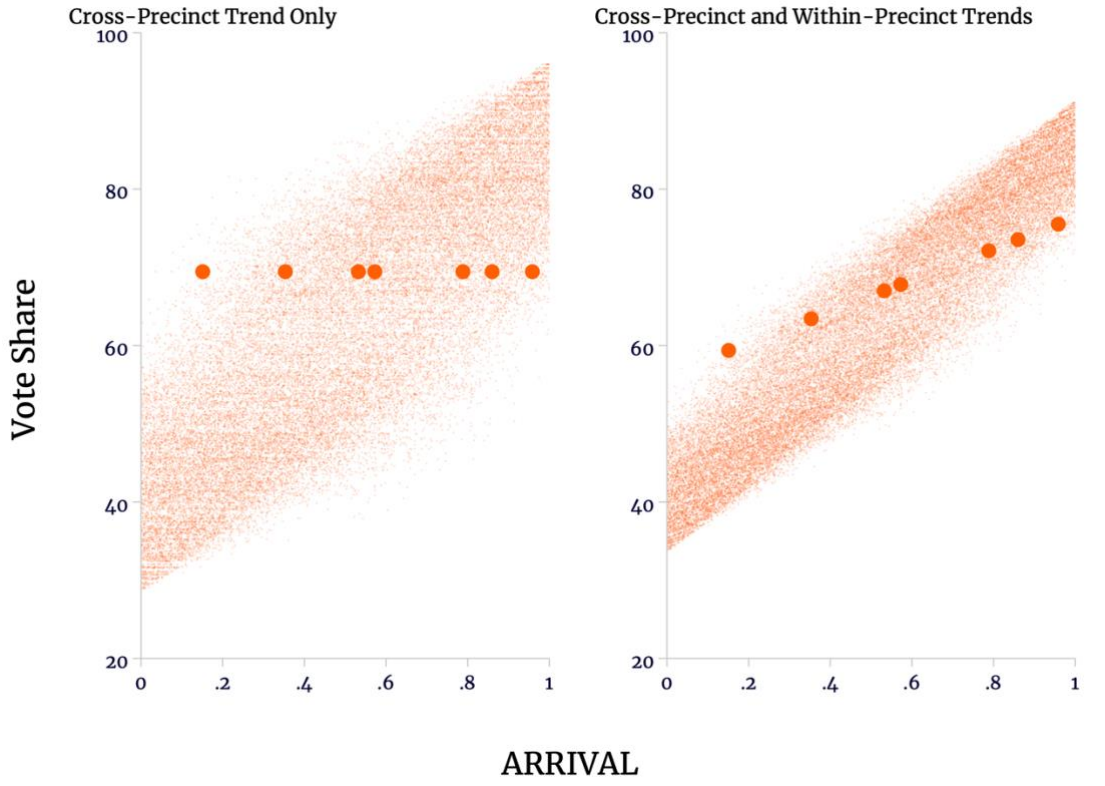
Going forward, we will, for clarity of illustration, generally avoid displaying the additional noise added in Figure 12. It is enough to recognize that there is considerable variation in the actual results from polling station to polling station, even if the voters are identical, statistically speaking.

In Figures 11 and 12, we produced data where geography has a direct impact on both vote shares and arrival order. In **Figure 13**, we contrast this with another election, with one critical difference: the inclusion of an additional trend in vote share directly determined by arrival order. That is, to some extent vote shares are higher the later they arrive — even within a precinct, and therefore with no regard for geography. Again, highlighting a single precinct helps to clarify the difference between the two election results.

On the left, we see that despite the overall trend toward higher vote shares, there is no such trend *within* the precinct. The overall pattern *across* precincts is driven by geography. On the right, there is a clear upward trend within the precinct and this is, therefore, not exclusively driven by the simulated geographic effects.

It is important to be able to make a distinction, statistically, between these two cases. The pattern on the left is only mildly surprising — *across* precincts those that reported late tended to report higher vote shares. On the right, higher vote shares are observed for late reporters *within* precincts as well.

Figure 13
Example Synthetic Data with Cross-Precinct and Within-Precinct Trends



Sources and Notes: Authors' calculations.

Neither is evidence of fraud *per se*. As described above, geographic effects may create the *nickels before dimes* pattern on the left. On the right, geographic effects contribute only partially to the overall trend. Still, there may be benign reasons for trends within precincts as well. Idrobo, et al., for example, posit that within precincts late reporters represent voters with surnames indicating lower socioeconomic status, driving increased support for Morales as well as late arrival due to tardiness or irregularity in processing the votes.[96] *Nickels before dimes*. This may not be a correct explanation, but this hypothesis is one of many that must be eliminated before fraud can be asserted. For purposes of our synthetic election data alone, we assume a given relationship between geography and vote preferences within an election that is independent of arrival. Thus, one way we may create fraud in the election is to add a within-precinct trend.[97]

96 Idrobo, Kronick, and Rodríguez (2022).
97 See Appendix C for more on differentiating different trends.

Applying Escobari and Hoover's Methods to Synthetic Data Reveals Flaws

We next go through the statistical methods presented in Escobari and Hoover's paper, applying them to synthetic data. In doing so, we see that to justify claims of fraud they require unreasonably strong assumptions about voting patterns.

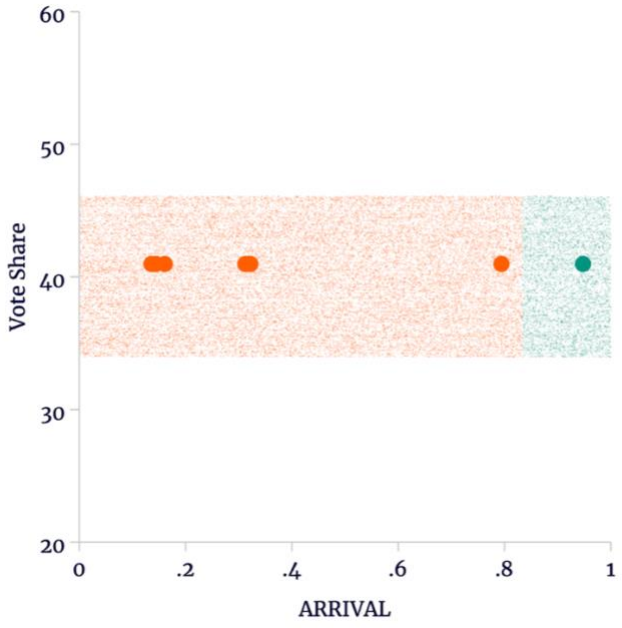
Difference Estimates

Suppose we have some election data and want to know if the late arrivals are in some way different from the early ones. One way we may test this informally is to simply inspect the data. In Figure 8, we show data with no association between geography and arrival order. The late arrivals (the last one-sixth of the results) are in green. There is no sign that the late ("SHUTDOWN") arrivals are different from the early arrivals.

Suppose instead that the SHUTDOWN tally sheets showed a different vote share. Using the same model, we might see something like **Figure 15**, where the late sheets have a slightly higher share than those that arrived earlier.

Figure 14

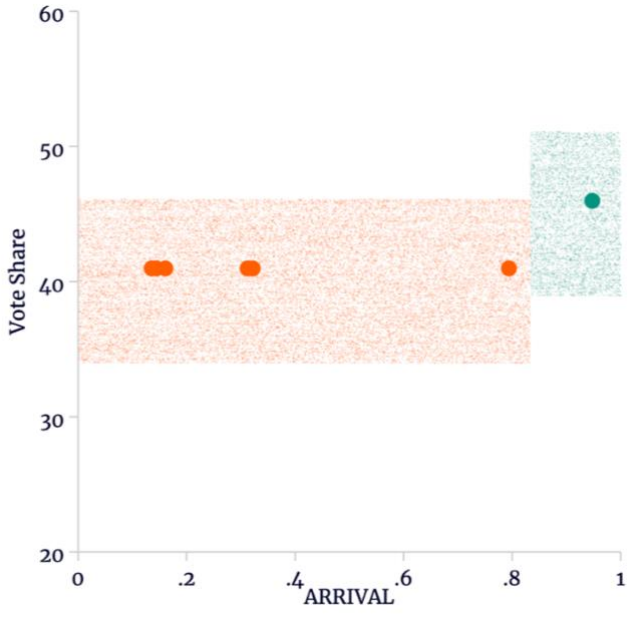
Synthetic Data with No Geographic Effect on Arrival



Sources and Notes: Authors' calculations.

Figure 15

Synthetic Data: Only Latest Arrivals Show Increased Support

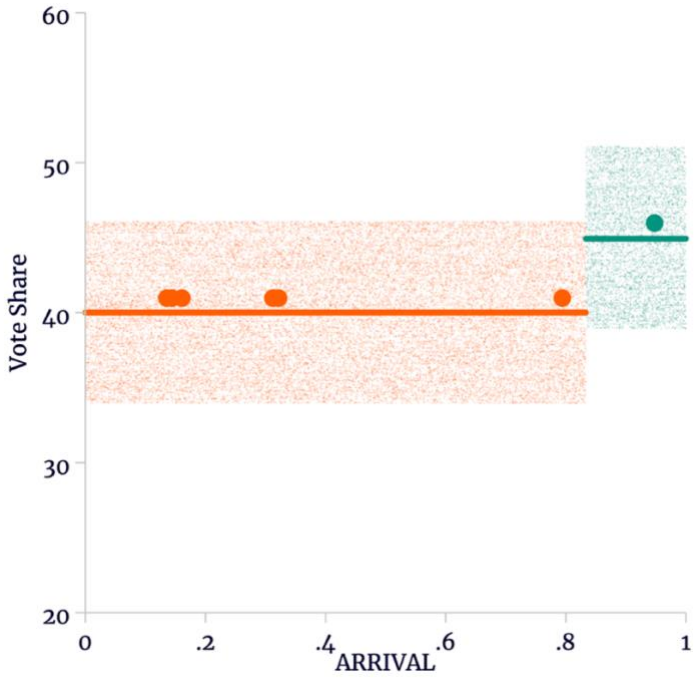


Sources and Notes: Authors' calculations.

While the data in Figure 15 suggests a difference within precincts, a sharp change may be detected even if vote shares are constant within each precinct. Suppose precincts transmitted all polling station data either all early or all late so that nickels (precincts entirely orange) come before dimes (precincts entirely green). Then, a benign shift in vote share may appear.

In either case, this sharp difference is the kind of effect difference models are good at measuring more formally. In **Figure 16**, we add to Figure 15 the expected vote share based on the difference model presented by Escobari and Hoover.[98]

Figure 16
Idealized Difference Model



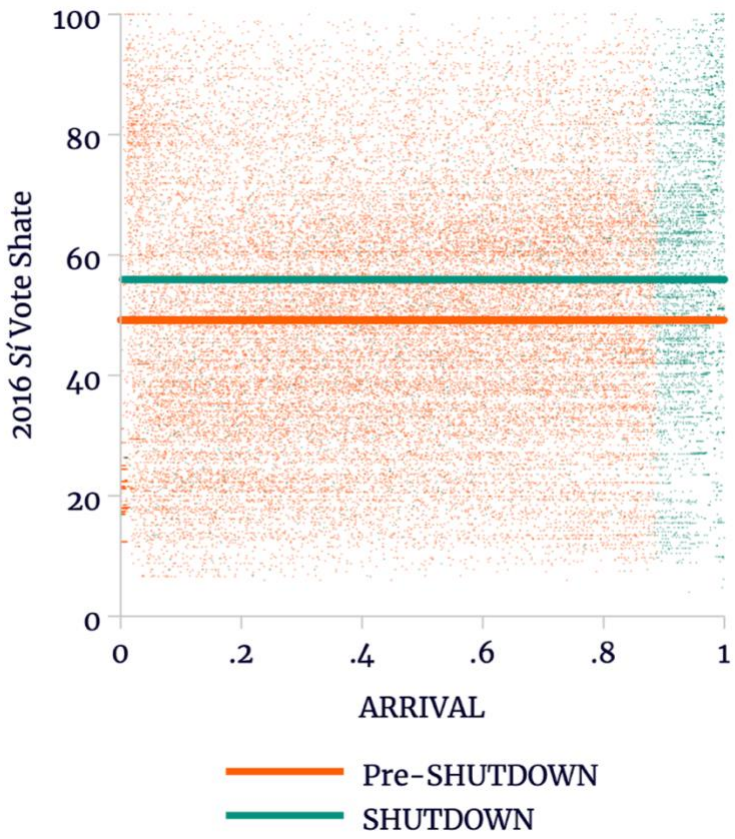
Sources and Notes: Authors' calculations.

If this were the actual data, we might have extra reason to be suspicious. In this synthetic data, 100 percent of the lowest vote shares came at the start of the count and 100 percent of the highest vote shares came at the end of the count. Perhaps there would be a reason for this within-precinct shift, perhaps not. However, the actual data is not so clean. So long as such a shift happens on average, the illustration is still appropriate. To see this, we may repeat the analysis on the actual data presented in **Figure 17**. The orange line reflects the average *Sí* share of the 2016 referendum matched to the polling stations included in the TSE announcement. The green line reflects the same for those stations reported after.

98 Escobari and Hoover (2020), Table 2, Column 2. See Appendix D for replication.

Figure 17

Difference Model on 2016 Results



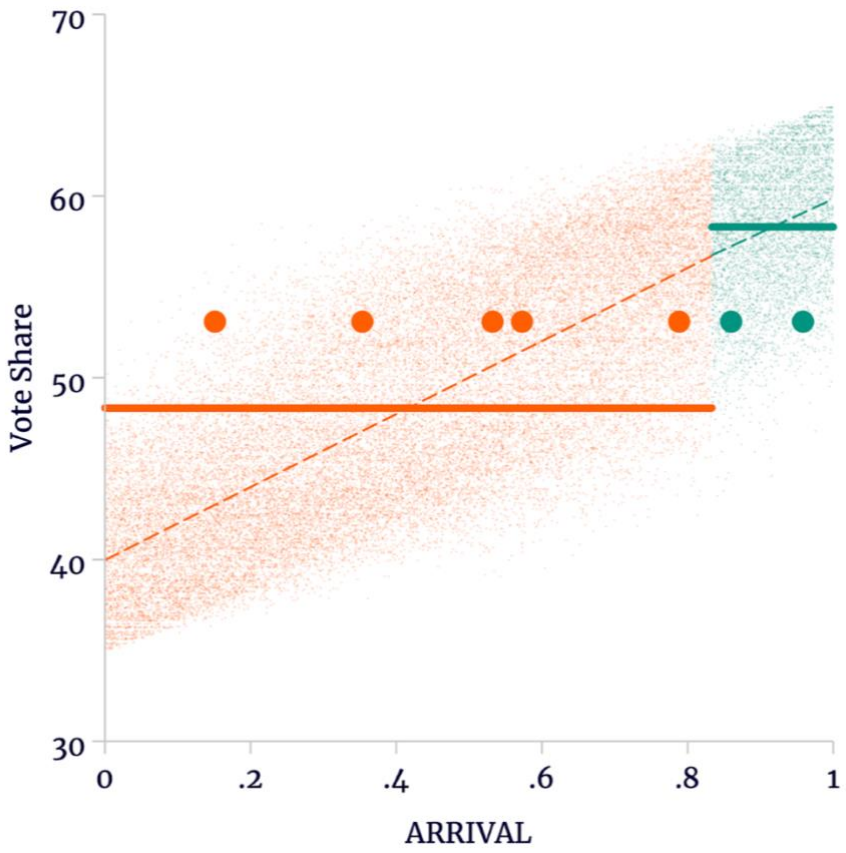
Sources and Notes: OEP (2016b), OEP (2016c), OEP (2017), OEP (2019b), OEP (2019c), OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations.

As discussed earlier at length, the post-announcement polling stations more heavily supported the referendum than the pre-interruption stations. This difference is also not necessarily indicative of fraud in the 2016 referendum. This model merely indicates that there is something different about the SHUTDOWN polling stations. This could be directly related to the fact that they were not included in the TSE announcement (which could be fraud) or a consequence of something not considered by the model resulting in the counting of nickels before dimes.

In **Figure 18**, rather than simply assuming the late arrivals have a lower share, we show synthetic data such that geography is associated with both arrival and vote share. This common determinant means that late arrivals will tend to be different — on average — than the early tally sheets. Though there is a considerable trend in margin as arrival times increase, there is no such trend within precincts. By failing to account for the trend, the model misleadingly detects a large jump in the last sixth of the data. Escobari and Hoover would characterize this jump as due to fraud, but no such jump exists in this data.

Figure 18

Difference Model Misleadingly Detects Unexpected Support in Late Arrivals



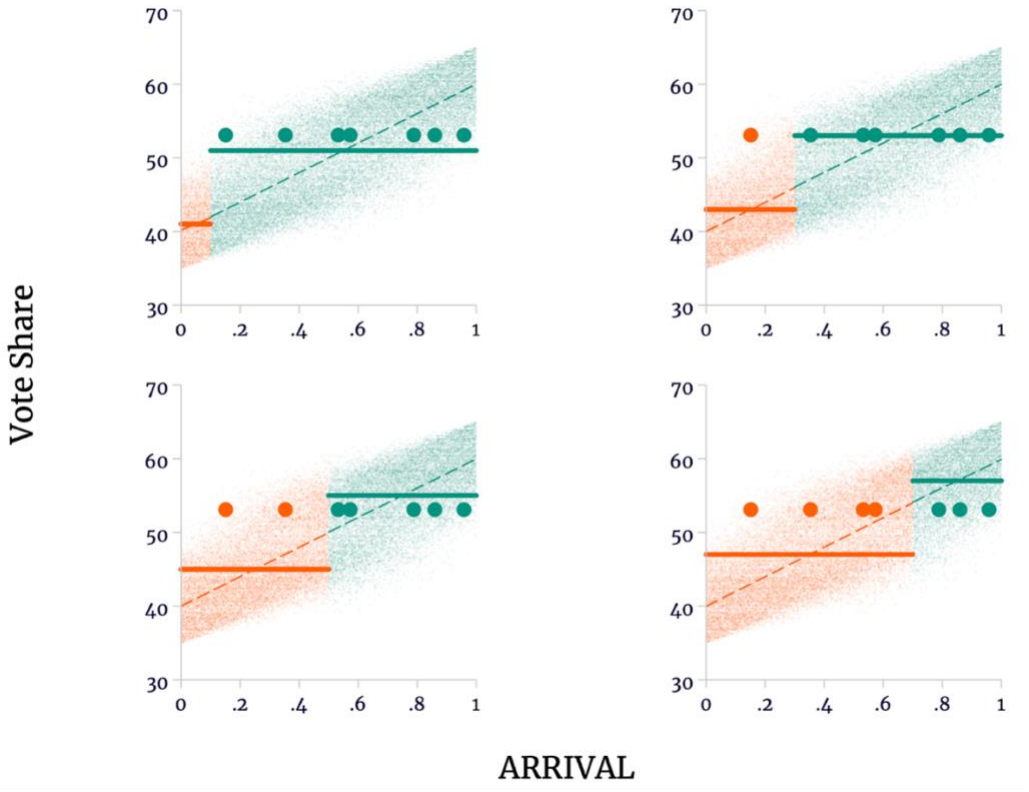
Sources and Notes: Authors' calculations.

We could construct data to match the left panel of Figure 6 more closely, with an upswing in support among those transmitting later. However, this would be unnecessarily complicated. The benign upward trend is sufficient here to see that the model, represented by the thick lines, is clearly wrong, overestimating the earliest arrivals and underestimating those just before the hypothesized break between early and late arrivals at the 5/6 mark. The thin dashed lines more closely represent the data. While the model correctly discerns that the late arrivals tend to have larger vote shares than the earlier arrivals, this has absolutely nothing to do, specifically, with the last sixth — the SHUTDOWN polling stations.

One way of seeing that the difference model does not inform us about the SHUTDOWN stations is to analyze other divisions of the synthetic data, as in **Figure 19**. Though the orange and green lines shift up as the hypothetical break in arrival increases, the difference between the two remains the same.

Figure 19

Difference Model with Varying Divisions



Sources and Notes: Authors' calculations.

More specifically, the difference in averages is exactly half the overall trend no matter where we set the mark.[99] The model tells us nothing about the SHUTDOWN polling stations, specifically. The detected change in vote share is due exclusively to an overall trend in vote share driven by geography alone and not at all caused by the late arrivals.

That there is an overall trend in the actual data is not in any dispute, as Escobari and Hoover affirm. If any part of this upward trend has a benign explanation, then the fraud claim resulting from the difference model is — generously — exaggerated. Nevertheless, Escobari and Hoover in their original paper insisted that this model, applied to actual data, suggested the *minimum* amount of detectable fraud.[100]

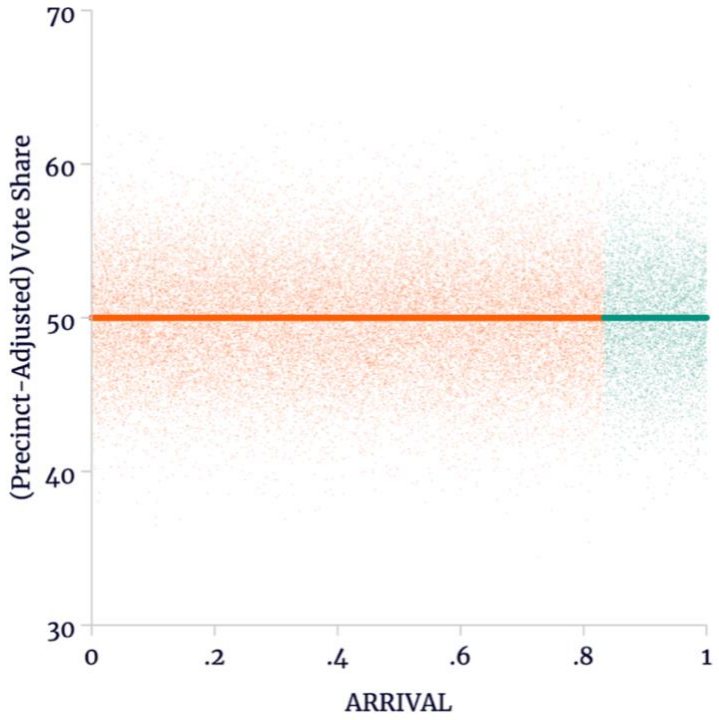
99 Because we construct the synthetic data to be linear and uniform in arrival time, this is true even if we choose a different break point. The average at arrival time 1 is 60, compared to 40 at arrival time 0. If we break at arrival q, then the average of the late shares is $60 - (60 - 40)(1 - q)/2$ and the average of the early shares is $40 + (60 - 40)q/2$. The difference is $(60 - 40) - (60 - 40)/2 = (60 - 40)/2 = 10$ no matter what q we choose.

100 Again, in their original paper, Escobari and Hoover referred to a model seeking to explain vote differences rather than differences in share — an additional problem.

To correct for geographic effects, Escobari and Hoover do report the results of adding precinct-level "fixed effects" to the model.[101] That is, they shift the shares in each precinct up or down so that each precinct has the same average (adjusted) share while retaining any differences among polling stations within each precinct. Because the polling stations in the synthetic data only differ (within a precinct) due to noise, this means there is no trend in the adjusted data at all.

Figure 20 shows the effect of adjustment on the model.[102] Given the adjustment, no increase in vote share for late arrivals is — correctly — detected.

Figure 20
Idealized Difference Model with Precinct-Level Effects



Sources and Notes: Authors' calculations

101 Escobari and Hoover (2020), Table 2, Columns 4–6 analyze margins (MAS-CC vote difference as a share of valid votes), rather than the shares (MAS or CC party votes as a share of valid votes) in Columns 1–2. We will continue to illustrate with shares in the synthetic data.

102 Figure 20 includes the added noise from random assignment of voters to polling stations. Without this, vote shares would be exactly constant within precincts, and so all vote shares would be adjusted to exactly the overall mean of 50. Likewise, we cannot present a similar figure for the 2016 data, as we match only at the precinct level. The precinct-adjusted data is by construction constant if the raw data is constant within precincts.

Furthermore, the core of the model makes sense applied to the adjusted data.[103] When applied to the official data, Escobari and Hoover report that geography accounts for about 98 percent of the observed change in vote margins. The remaining difference (fewer than 3,600 votes out of more than 6 million) is an order of magnitude too small to account for Morales's 10.56 percentage point victory.[104]

Again, attributing even this remaining difference to fraud requires the assumption that it is the late tally sheets that are in question as opposed to the early tally sheets — a Panic Pete problem. The model itself does not distinguish between an unexpected gain of 3,600 votes for Morales in the SHUTDOWN polling stations from an equally valid alternative: an unexpected loss of 18,000 Morales votes in the TSE announcement.[105]

If an excess of 3,600 votes were otherwise proved to be the product of a fraud perpetrated by Morales, then the former president ought to face legal consequences. But even accepting that these votes were somehow invalid, that leaves Morales with more than 10.5 percentage points legally cast votes than for Mesa, and a corresponding first-round victory.

However, Escobari and Hoover in their original paper write that with the use of these precinct-level fixed effects:

> [Variation due to] Electoral fraud will be erased if, for example, electoral fraud affects all polling stations within the same precinct. In this case we will have that fraud is perfectly collinear with the fixed effect, hence [the fixed effect] will wipe out some fraud and we will not detect it. Moreover, identification will be coming only from precincts that have polling stations before and after the shutdown.[106]

103 The vote margin is the difference in vote shares for the top two candidates. Columns 3–6 of their Table 2 apply models to the vote margin (rather than share) for a particular party. For purposes of illustration, it doesn't matter if we relabel the synthetic vote shares as margins.

104 Further, the result fails to reject the hypothesis of zero "fraud" once vote shares are correctly computed. See Rosnick (2019), Table 5.

105 Escobari and Hoover (2019) say the late votes are unexpectedly high in support of Morales, but this is not statistically distinct from saying the early votes are unexpectedly low in their support for Morales. We may resolve the difference by making a (dubious) inference that the unexpected difference (either way) is due to fraud. If we reduce the late support for Morales to bring the margin in line with the early, this comes to a loss of about 3,600 votes relative to the official tally. If we increase the early support to bring the margin in line with the late, this comes to a gain of 18,000 (about -5 x 3,600) because the difference in margin is applied to about 5 times as many votes.

106 Escobari and Hoover (2019), 5.

In other words, if a precinct is entirely reported late and has unusually high support for Morales, then that will be captured by the precinct fixed effect and not SHUTDOWN — the variable they use as a proxy for fraud. The use of polling-station level effects in this model thereby eliminates the positive cross-precinct trend, leaving much less of the pre-post announcement differences to be explained. That is, with polling station effects, the SHUTDOWN effect measures only the within-precinct differences in margins between late and early stations.

If the unusually high support in the late precinct is due to fraud, then SHUTDOWN will fail to capture it. Having possibly failed to capture some potentially fraudulent votes, Escobari and Hoover feature the large unadjusted result (like Figure 18) as a minimum amount of "fraud" and dismiss their own precinct-adjusted result (like Figure 20) as indicating an impossibly small effect.

However, if the unusually high support in the precinct is explicable by any factor other than fraud, then the unadjusted SHUTDOWN effect incorporates more than fraud. In effect, Escobari and Hoover simply attribute the entire cross-precinct trend to fraud.

As we observed above, there is every reason to believe that the SHUTDOWN polling stations — those not included in the TSE announcement — were much, much more favorable to Morales and that Escobari and Hoover's estimate of fraud is utterly implausible.

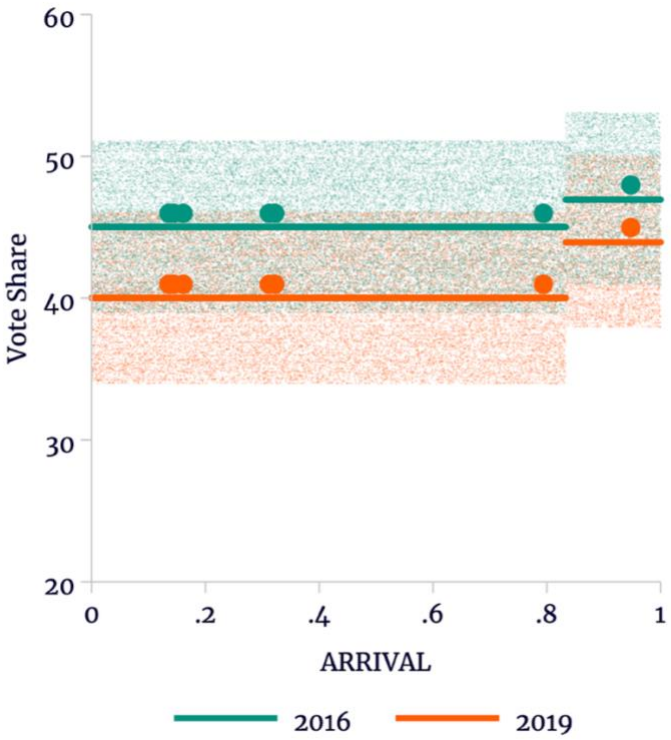
Escobari and Hoover agree in their updated paper that the late reporting stations clearly favored Morales in 2016. It would appear implausible, then, that their models could put the outcome of the election in any doubt. Still, Escobari and Hoover proceed to offer inadequate analyses employing the 2016 data.

Difference-in-Difference Estimates

Escobari and Hoover next expanded their model to see if the increase in Morales's support went beyond what could be expected by taking account of the referendum results in 2016. To do this, they estimated four averages: the early and late shares for 2016, and the early and late shares for 2019. They reason that if the increase from early to late in 2019 exceeded the increase from early to late in 2016, then this excess increase indicates fraud. That is, in an idealized case with no trends, we might see something like **Figure 21** where the green 2016 shares were higher in the last sixth. Even though the late stations in 2019 supported the incumbent less than the early

stations in 2016, the difference between 2016 and 2019 was narrower among late stations than early.[107]

Figure 21
Idealized Difference-in-Difference Model



Sources and Notes: Authors' calculations.

This is called a "difference-in-difference" model because we are comparing the changes over time and then across groups. We take the difference in early and late 2019 (44 - 40 = 4) and subtract the difference in early and late 2016 (47 - 45 = 2) to find that the 2019 shares rose by an extra 2 percentage points. (Equivalently, we may take the late difference of 44 - 47 = -3 percentage points between 2019 and 2016 and subtract the early difference of 40 - 45 = -5 percentage points between 2019 and 2016 for the same 2 percentage point "double-difference.")

Again, such an analysis simply observes that though the late arrivals showed additional support in both elections, the additional support was greater in 2019. Fraud presupposes that the 2 percentage point double-difference is suspect. The analysis does not suggest the late 2019

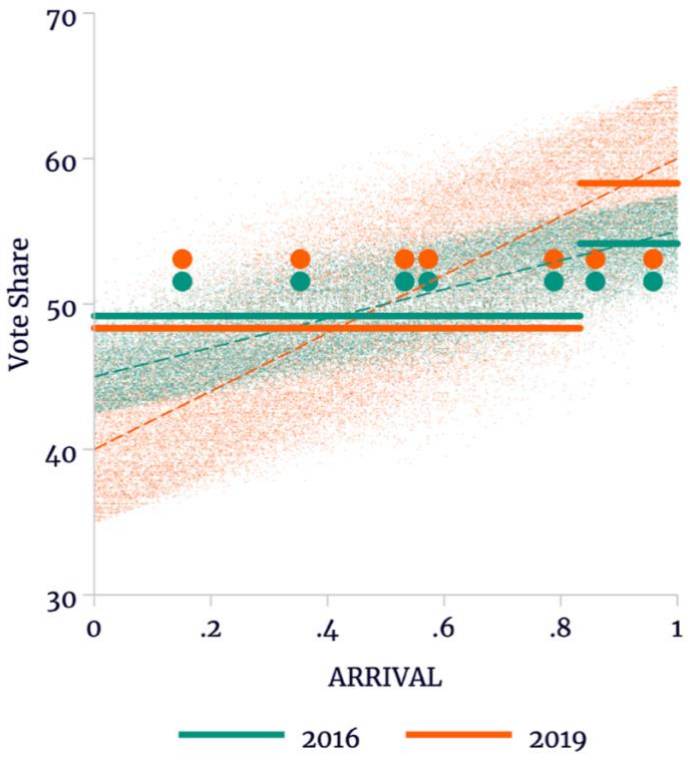
107 Note again that results in Figure 16 come in pairs — every 2019 result in orange corresponds to a 2016 result in green — and these pairs "arrive" simultaneously. Escobari and Hoover claim to have arranged their data in this fashion.

shares are abnormally high rather than the lower early shares in 2019, or the late shares in 2016. Fraud in late 2019 arrivals is a possible conclusion if we assume that:

1. the 2016 and early 2019 votes are free of fraud, and;
2. there is no benign explanation for the difference-in-difference of 2 percentage points.

Just as in Figure 18, the model may still report differences when there are stable trends underlying the data. In Figure 18, the model failed because arriving tally sheets showed a general trend toward increasing shares (trending upward when compared to a flat baseline). Below, in **Figure 22**, the double-difference model fails when the 2019 slope is different from the 2016 slope.[108] While this could be due to an inexplicable trend over arrival times, here the 2019 results are simply more sensitive to geography than the 2016 results — opposition polling stations increased in their opposition and incumbent polling stations increased in their support. The model again proposes a sudden shift in the last sixth of the data that does not actually exist.

Figure 22
The Difference-in-Difference Model Fails Where Trends Differ



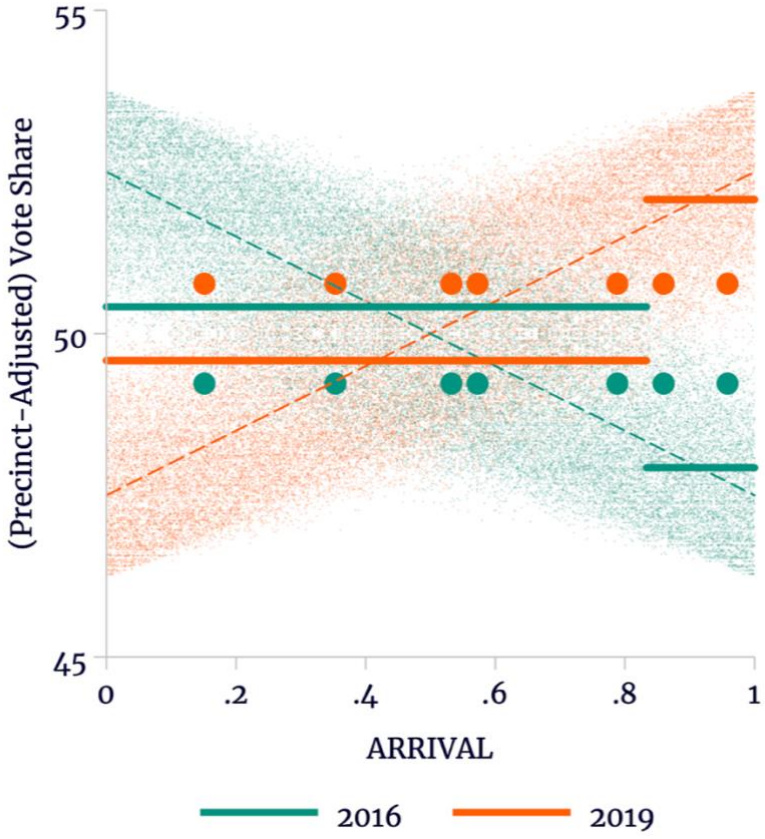
Sources and Notes: Authors' calculations.

108 The difference model is in effect a special case of the difference-in-difference model: one where the 2016 slope is assumed to be zero. In the difference model, a slope greater than zero drove the model to find a change in support. In the difference-in-difference model, a slope greater than that observed in 2016 does likewise.

In this case, the double-difference comes to 5 percentage points. However, the difference is entirely across precincts; late arrivals within each precinct have no higher share than the early arrivals. Nor is there any late break in trend whatsoever. The difference-in-difference is because voters at individual *precincts* were more consistently unified in 2019.

In **Figure 23**, we see a slightly tweaked model.[109] The difference from Figure 22 is the addition of precinct-level effects. That is, vote shares are adjusted in the model so that the average vote share in any given precinct — across all polling stations and across both years — is the same for every precinct.

Figure 23
Precinct Effects Do Not Improve the Difference-In-Difference Model



Sources and Notes: Authors' calculations.

In the difference model (see Figure 20), the polling-station plus noise explained all the available data, so the adjusted vote shares wound up as a constant plus noise. With the addition of 2016 data, this is no longer true. Early arrivals tend to have higher vote shares in 2016 than in 2019,

109 Escobari and Hoover (2020), Table 3, Column 6.

and vice versa for late. So now the precinct average of 2016 and 2019 are constant, and the overall upward trend is removed, just as in the difference model.

Note that the underlying data is the same as in Figure 23 — only the model changes.

Escobari and Hoover write that this model "controls for heterogeneity across precincts and polling stations that did not change between 2016 and 2019."[110] To be more precise, the effects control for *voting tendencies* that vary across precincts, but the effects do not control for voting tendencies that vary across elections.

By way of example, suppose that the campaigns sow divisions between rural and urban voters, so that urban voters move toward the opposition and rural voters increase support for the incumbent. If there is a bias toward counting urban votes first, then contrasted with 2016 the incumbent vote share in 2019 will be relatively low early and relatively high late. The heterogeneity — the geographic variation across precincts — does not change between 2016 and 2019. However, the effect of that heterogeneity on voting tendencies is not stable across elections, and Escobari and Hoover do not control for it.

Likewise, it does not control for anything like a precinct's soft support (illustrated in Figure 4) that is itself unchanged from one election to another. While the soft support in our example does not change across elections, it swings from support for the incumbent to support for the opposition. Escobari and Hoover implicitly assume that soft support is constant across polling stations.

Because the precinct-level effects pick up the overall trend inclusive of both elections, adjusted vote shares for each year trend in equal and opposite directions; the difference in slopes is maintained from Figure 22 to Figure 23. The difference-in-difference (half the change in slope) is therefore the same. The model says that in 2019 the late shares rose by 2.5 percentage points, while those in 2016 fell by 2.5 percentage points — leaving the same 5 percentage point double-difference observed in the basic difference-in-difference model.

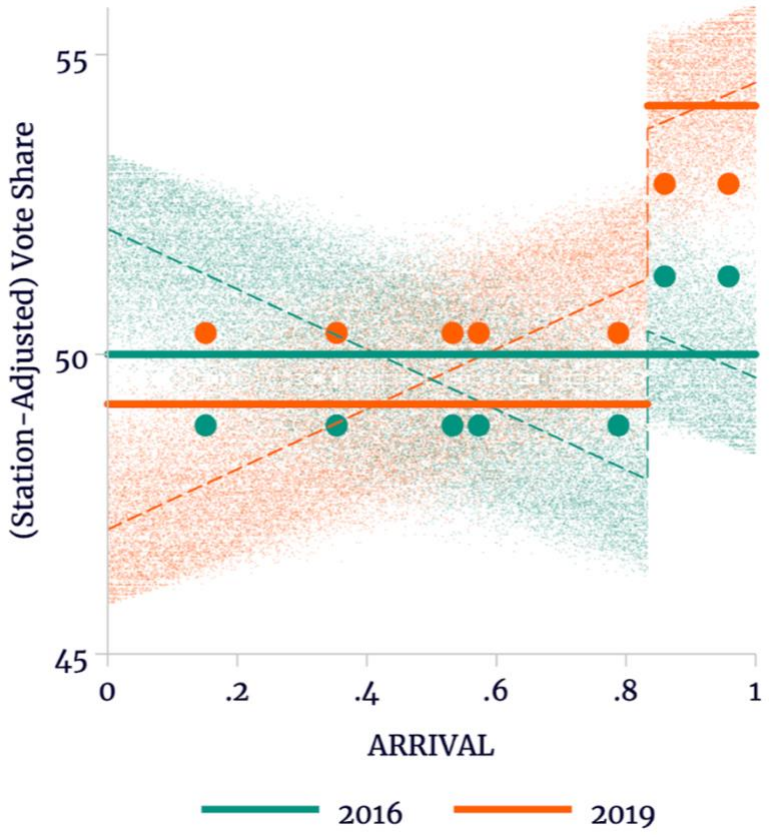
We could tweak the model again, following Escobari and Hoover and adding polling-station fixed effects in lieu of precinct-level.[111] Here, things get stranger. Because vote shares do not vary apart from noise in the synthetic data, this would seem no different than using precinct-level effects.

110 Escobari and Hoover (2020), 13.
111 Escobari and Hoover (2020), Table 3, Column 4.

However, polling stations are defined as late based on 2019 arrival alone, so there is no way of distinguishing (*a*) a late polling station that has an unexpected vote share because it is late, from (*b*) a late polling station that has an unexpected vote share because it has an unexpected polling-station effect.

To identify the differences between 2016 and 2019, Escobari and Hoover assume that in 2016 late polling stations have on average the same *station-adjusted* vote share as early polling stations. That is, though the solid-green model predictions on the adjusted data of Figure 23 shifts down for the last sixth of the data, there is no such shift in **Figure 24**. But the difference in vote shares from 2016 to 2019 within each polling station must be preserved, so all the late vote shares for both years must be adjusted upward (and all early vote shares adjusted downward), creating breaks in trends for both the 2016 and 2019 vote shares.

Figure 24
Escobari and Hoover's Polling Station Effects Create Breaks in Trends

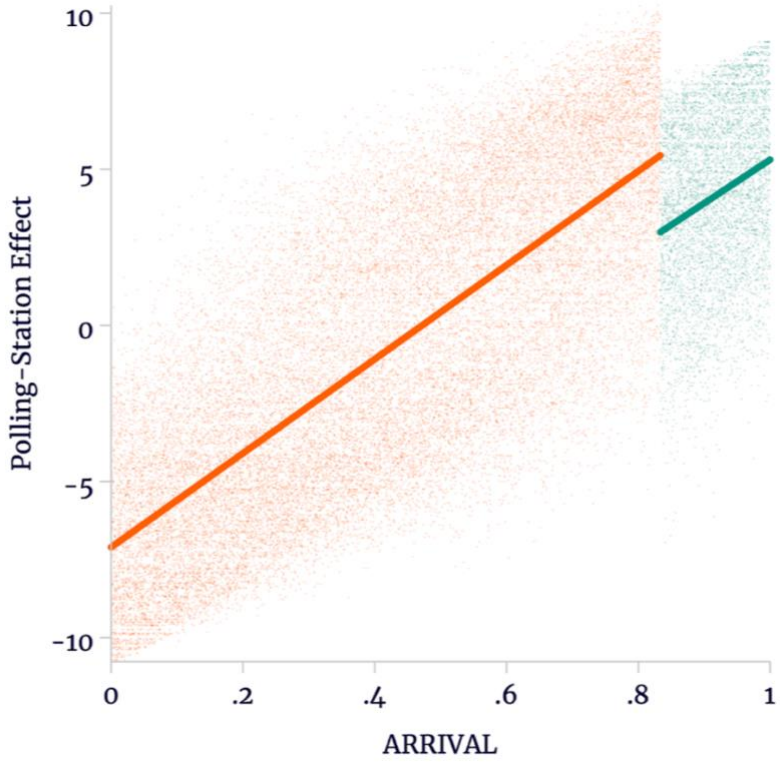


Sources and Notes: Authors' calculations.

It is imperative to recognize that the breaks in trend observed in the adjusted data are completely manufactured by Escobari and Hoover's modeling choice. We see in Figure 24 that within the highlighted precinct the adjusted vote shares jump upward even though the actual vote shares are constant throughout the precinct.

Equivalently, the late polling-station effects (which have been subtracted from the actual vote shares to create the adjusted vote shares) must shift suddenly downward, as in **Figure 25**. Escobari and Hoover have implicitly assumed that there is something inherent in the late stations that makes them less favorable to Morales than the trend suggests. Again, we know this is not true in the synthetic data.

Figure 25
Escobari and Hoover's Polling Station Effects Drop Suddenly



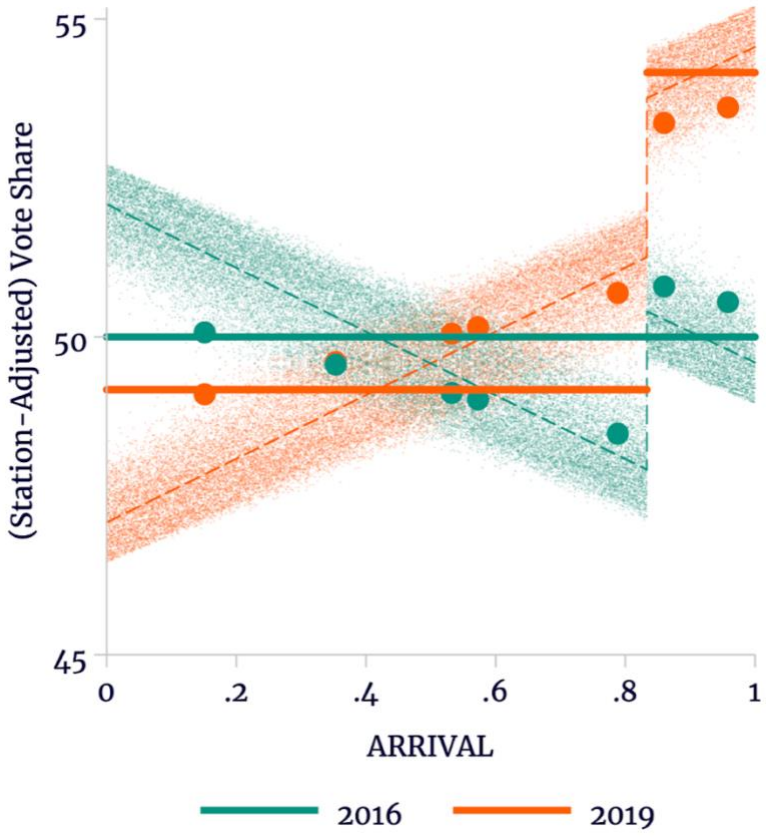
Sources and Notes: Authors' calculations.

Though it may seem very odd, this fortunately all comes out in the wash because it is the double-difference that is of interest. So long as the effects for polling stations are all shifted by the same amount in each group (early and late), the differences between 2019 and 2016 for each group are unchanged and therefore the double-difference (difference in each group's difference) is not changed. Because there are no within-precinct differences, polling-station level effects can add nothing to the analysis.

In **Figure 26**, however, we repeat the analysis using different synthetic data. This time the 2019 results include an additional within-precinct trend.[112] The overall difference in trends is still 10 percentage points of vote share, but now half is due to geography and half is artificial. Yet the main result is still the same: a double-difference of 5 percentage points. The model does not distinguish trends due to geography from trends within precincts.

Figure 26

Same Overall Difference in Trends Creates Same Double-Differences Whether or Not Any Within-Precinct Trend Exists



Sources and Notes: Authors' calculations.

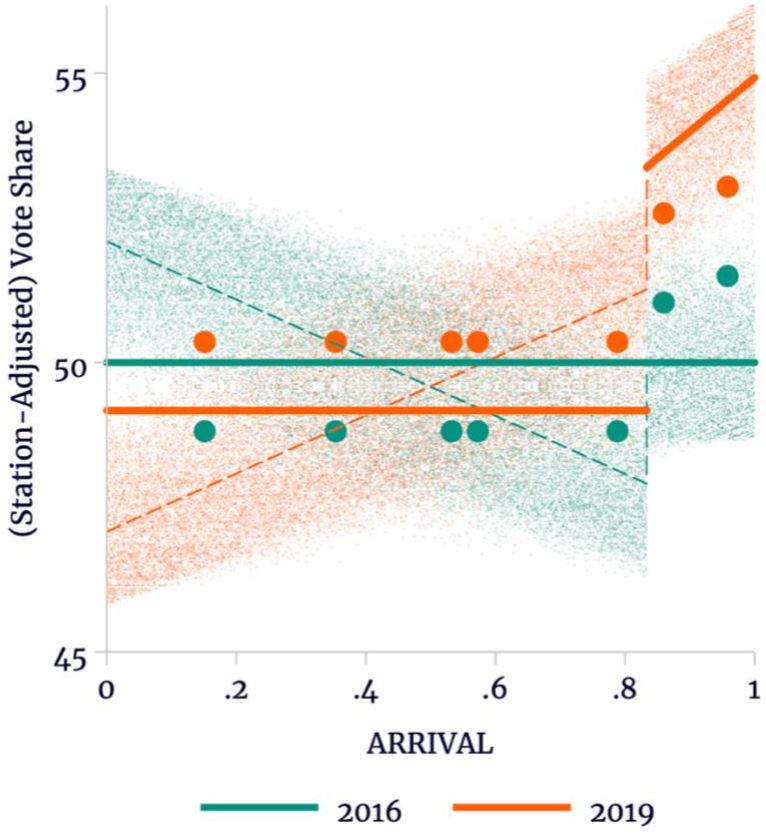
Put simply, these models do nothing to explain why the vote shares in 2019 grow faster than those in 2016, but rather pick up the difference in trends and report it as a shift in the late arrivals. Escobari and Hoover simply take this indirect measurement of the observed difference and call it fraud.

112 Note that the within-precinct trend is visible in the adjusted data. Even among the early stations, the adjusted 2019 results trend up and the 2016 trend down.

"Evidence" of Widespread Fraud

Escobari and Hoover proceed next to permit a benign common trend as well as precinct effects.[113] This adds absolutely nothing, as it is the difference in trends that drives the observed double-difference.[114] Again, Escobari and Hoover pick up the fact that 2019 results are more geographically sensitive than in 2016.

Figure 27
An Additional Trend for Late Stations in 2019 Doesn't Fix the Double-Difference



Sources and Notes: Authors' calculations.

Their next model is a step backward.[115] There, Escobari and Hoover again add polling-station level effects. Again, facing a choice, they simply impose on the adjusted model that there be no

113 Escobari and Hoover (2020), Table 4, Column 3.
114 The actual addition here is the irrelevant "COMPUTO" variable: the timing of the polling station in the official count.
115 Escobari and Hoover (2020), Table 4, Column 4.

2016 trend. The polling-station effects must compensate for this, resulting in the adjusted data and model of **Figure 27**. Again, Escobari and Hoover's model tries to accommodate the difference in trends, and picks up a nonexistent trend in the late polling stations as well as a nonexistent break.

This is almost identical to Figure 24, except that it actually picks up the post-interruption trend: the orange line slopes upward for late arrivals in 2019 only. However, the *adjusted* vote shares trend upward in both years within-precinct. Escobari and Hoover highlight this difference, but it is, again, completely artificial. The (now varying) double difference is still misleading for the same reason as before. The apparent break is due to the difference in trends.

At long last, Escobari and Hoover acknowledge what has been clear, writing that the "gap [between 2016 and 2019] appears to be closing throughout" arrival times.[116] As we have noted above, there may be perfectly benign reasons for this. The existence of soft support, or a relative decline in urban support, may cause an early gap to shrink. The authors are undeterred, however, writing:

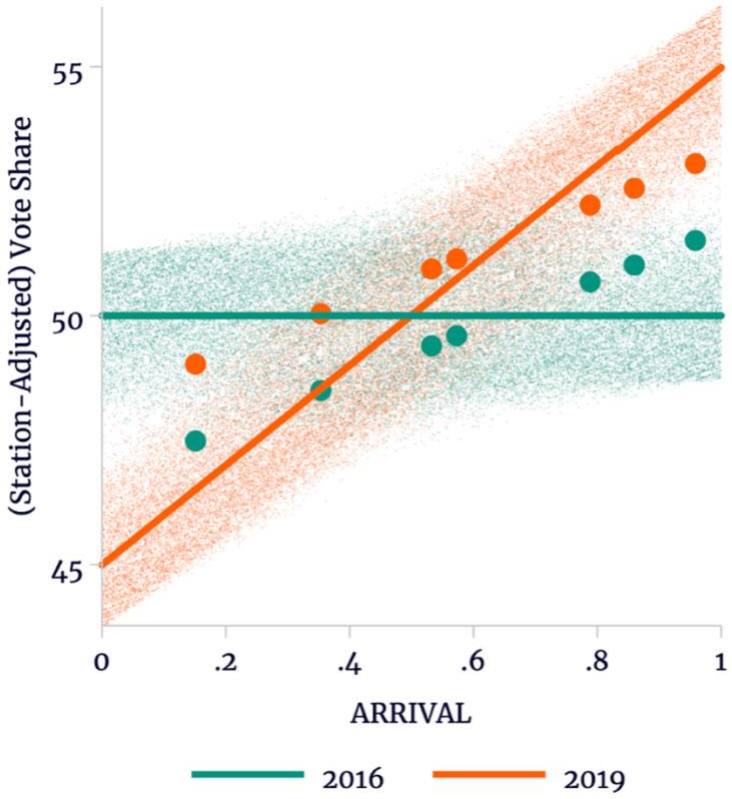
> We now turn to assess the hypothesis that fraud existed in pre-shutdown ballot boxes. We do so by relaxing the parallel trends assumption and by disentangling the effect of timing of the arrival from the effect of the shutdown.[117]

The new underlying assumption is again that there is no 2016 trend in the adjusted data, but they now allow the model to pick up the difference in trends, as in **Figure 28**. This is precisely what we would want from a model, if the difference in trends would be linear absent any fraud. This model correctly discerns that there is no break in the last sixth of the synthetic data.

116 Escobari and Hoover (2020), 18.
117 Escobari and Hoover (2020), 18.

Figure 28

Adding an Overall Trend for 2019 Picks Up Differences in Cross-Precinct Trends as Well as Within-Precinct



Sources and Notes: Authors' calculations.

In the synthetic data, the trends are explicitly linear, so there is no break — no unexpected difference in the last sixth of the polling — and so no unexpected double-difference.

Escobari and Hoover's reported results indicate that taking the difference in slopes into account, the model fails to explain only 1 percentage point of Morales's margin at the late polling stations. This comes to about 0.18 percent of all valid votes. Even if Escobari and Hoover attributed this unexplained double-difference to fraud, it would be about one-third the amount necessary to reverse the official outcome of the election.

This ought to be the end of the story, but Escobari and Hoover offer a completely different interpretation of their results. Rather than "relaxing the parallel trends assumption" by allowing the differences from 2016 to 2019 to trend, they argue that the differences in slopes is itself evidence of fraud. Because the effect is 5.72 percentage points over the course of arrivals, they infer half this amount — 2.86 percent of all valid votes — to represent fraud in favor of Morales.

It is hard to overstate how pernicious this shift in focus really is. Their work to this point has expressly centered on the interruption of public reporting of the unofficial results after the TSE announcement, arguing that the interruption casts suspicion on polling stations reported after that point. They even argued that the fact that the post-interruption results were (to them) seemingly scattered at random throughout the official count served as further evidence that the electoral authorities had attempted to evade detection.[118] However, the fact that the results associate with arrival time — and ultimately that the parallel trends assumption is false — invalidates all their analyses so far.

Even though there may be benign explanations for nonparallel trends, and Escobari and Hoover recognize the problem and propose a partial fix, they simply argue that if their central assumption of parallel trends is wrong (thereby invalidating all their previous analyses) then *that too* is evidence of fraud.

Even though there may be benign explanations for nonparallel trends, and Escobari and Hoover recognize the problem and propose a partial fix, they simply argue that if their central assumption of parallel trends is wrong (thereby invalidating all their previous analyses) then *that too* is evidence of fraud.

They do this by characterizing the difference in slopes (the orange line in Figure 28) as a 2019-specific within-precinct trend, which is clearly untrue in the synthetic data. Within precincts, the differences between 2016 and 2019 synthetic vote shares are — by construction — constant. There can be no difference in within-precinct trends. We see in Figure 28 the *adjusted* margins in the highlighted precinct trending up, but exactly in parallel. Again, there is no 2019-specific trend within precincts. The interpretation offered by Escobari and Hoover is wrong for the synthetic data and there is no reason to think it any less wrong when applied to the official results in Bolivia.

Even if we assumed that the overall trend is indicative of fraud, there is absolutely nothing that points to it as fraud *on behalf of the incumbent*. It is equally valid to argue that the difference in trends is consistent with decreasing fraud as the opposition finds fewer and fewer opportunities to manipulate the count.

This is exactly the Panic Pete problem we observed in **Figure 5**. The difference in slopes — even if simply assumed to be the result of fraud — simply doesn't tell us whom the fraud favors.

118 Escobari and Hoover (2019), 6. This argument does not hold up to scrutiny. See Rosnick (2019).

Moreover, it is only the strong, unjustified, assumption that the trends should be parallel that even suggests any possibility of fraud.

Worse, Escobari and Hoover erroneously report their observed effect on the official data (0.57 percentage points of margin per 10 percent of the count) as consistent with Idrobo *et al.* who themselves reported a within-precinct trend of about 0.5 percentage points over 100 percent of the count — a full order of magnitude smaller than the effect claimed by Escobari and Hoover.

Escobari and Hoover do put forth a theory of fraud that they say could explain the trend. Specifically, they suggest:

> It takes time to implement fraud, so polling stations that took longer to arrive at the Electoral Court are more likely to be contaminated with fraud. Hence ARRIVAL captures a treatment that grows gradually ...[119]

To be clear, they offer no explanation for why later polling stations would have more fraud than earlier ones. It is the prevalence of the fraud that may increase over time, not the size of the fraud at any given station. They suggest time is required — "For example, to rewrite the booth minutes and forge signatures"[120] — but *rewriting* implies that there was a first tally for which none of the participants retained their copy, and none waited for the actual transmission of the results. Barring that, it would be faster — not slower — to simply fabricate totals than to count the actual ballots. A Potemkin count to avoid suspicion at nearby polling stations in the same precinct would be at least as rapid. The mechanism is implausible, and is indistinguishable from the completely benign explanation that vote shares were more geographically sensitive in 2019 than in 2016 — that support was particularly low on average, early, for Morales, where his support was already relatively weak and fleeting. On average, areas where Morales's support was strong and unyielding were counted later. *Nickels before dimes.*

Finally, in discussing these results, Escobari and Hoover repeat unsubstantiated suspicions regarding the vote from Argentina, citing a string of false and misleading statistics.[121]

119 Escobari and Hoover (2020), 20.
120 Escobari and Hoover (2020), 20.
121 For details, see Appendix E.

Triple Differences

Escobari and Hoover next stretch even further, with "triple differences."

Their difference models contrast early and late vote shares. Their difference-in-difference models contrast early and late vote shares as they differ from 2016. As observed above, the difference model is a special case of difference-in-difference, where the 2016 vote shares are implicit and do not change from early to late arrivals. The contrasting difference is assumed to be zero.

A "triple difference" model extends this again. Escobari and Hoover seek to contrast the observed double-difference (2016 vs. 2019, early vs. late) to a second double-difference. That is, the difference-in-difference model is a special case of the triple-difference, where the contrasting difference-in-difference is assumed to be zero.

The idea is that even if, say, the 2019 trend in MAS vote share differs from the 2016 trend in MAS vote share, it may be that the same difference in trends shows up for non-MAS parties as well. In Escobari and Hoover's interpretation, to the extent any difference in trend is not confined to MAS, it is not fraudulent (or at least not a MAS-specific fraud).

Thus, if the MAS share rises 10 percentage points faster in 2019 relative to 2016, resulting in a double-difference of 5 percentage points, but the CC share rises 6 percentage points slower than 2016 for a double-difference of -3, then the triple-difference is 5 - (-3) = 8 percentage points.

Of course, we don't expect MAS shares to run in parallel with CC. Nor would we expect anything causing MAS shares to rise more rapidly to also cause CC shares to rise more rapidly. To perform this analysis, Escobari and Hoover must pick parties that they can justify, assuming such parallel trends (and to *Sí* and *No* in 2016). They pick Movimiento Tercer Sistema (MTS) and Bolivia Dice No (21F), respectively, to parallel MAS and CC. There are two very important problems with this approach.

First, Escobari and Hoover misstate the formula for a true triple-difference.[122] There should be eight terms in a triple-difference (seven stated explicitly, and one constant implicit in the set of controls).[123] In the case of Escobari and Hoover's triple-difference, the double-difference for MAS

122 Escobari and Hoover (2020). Equation 4 shows six terms, not eight.
123 See, for example, Escobari and Hoover (2020), Equation 5.

in 2019 is made in comparison to *Sí* in 2016,[124] but the double-difference for MTS in 2019 is also made in comparison to *Sí* in 2016. When the triple difference is calculated, everything involving 2016 simply drops out. Thus, we are left with not a true triple-difference but a double difference with parties (comparing MAS to MTS) instead of years (2019 to 2016). This is unclear in their formulation because two terms are missing. These missing two cancel exactly two terms written, leaving only four terms — not six or eight. A correct construction proceeds as follows, starting with single differences and working up:

$$\beta^{D}_{MAS} = LateMAS - EarlyMAS$$
$$\beta^{D}_{MTS} = LateMTS - EarlyMTS$$
$$\beta^{D}_{YES} = LateYES - EarlyYES$$

$$\beta^{DD}_{MAS} = \beta^{D}_{MAS} - \beta^{D}_{YES}$$
$$\beta^{DD}_{MTS} = \beta^{D}_{MTS} - \beta^{D}_{YES}$$

$$\beta^{DDD} = \beta^{DD}_{MAS} - \beta^{DD}_{MTS} = \beta^{D}_{MAS} - \beta^{D}_{MTS}$$

The flip side is that two of the terms in their regression equation (e.g., α_3 and , δ_2) must be zero because there is no difference between what they call "TREAT" and non-"TREAT" results in 2016 — both are *Sí*. "Treatment" in the sense suggested by Escobari and Hoover can only have an effect in 2019, where we can distinguish MAS and MTS votes. Restricting the model to a double-difference using only 2019 data has absolutely zero impact on the estimated "triple" difference. Replacing actual 2016 results with completely random numbers has no impact on the triple difference.[125]

The other problem is that the parallel trends assumption makes no sense in this context. Escobari and Hoover assume parallel trends based on the correlation between MAS and MTS *votes* in the early results. However, this is not sufficient for multiple reasons.

First, the baseline differences are assumed equal, even though MAS is a major party and MTS minor. Suppose voters in one precinct are twice as likely to vote for MAS as are voters in another

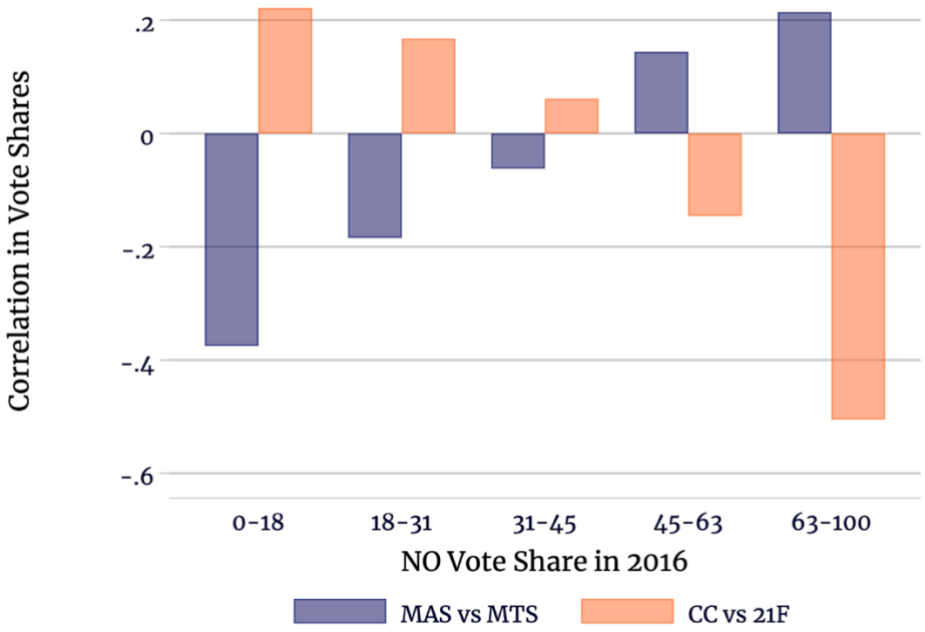
124 More precisely, they look at differences between MAS and CC and differences between *Sí* and *No* — margins, rather than vote shares. Just as with the previous analyses, the problems carry over. We use shares rather than margins here for clarity.

125 The number of observations is even more confusing than in earlier results. Escobari and Hoover (2020) report 136,318 observations, though there ought to be at most 2 per polling station pooled across years — or 129,844. Prior to their Table 6, they report no more than 66,595 observations (Table 3, Column 4) — a number inexplicably more than the total number of polling stations available in the data. Twice this is 133,190, still short by 3,128. Consequently, there is no sensible way to replicate the results.

precinct — say, 20 percent compared to only 10. Are voters also twice as likely to vote for MTS — say, 10 percent compared to only 5? Then the MAS vote share rises twice as fast as that of MTS; they are not parallel. We can get the shares to follow parallel trends by tripling, rather than doubling, the MTS share of the vote.

Second, Escobari and Hoover justify assuming parallel trends based on the (pre-interruption) correlation between votes for MAS and votes for MTS. The vote shares do correlate. However, they are complements in some geographic areas (parties succeeding and failing together), and in others they are substitutes (approaching zero-sum competition for votes). In addition, Escobari and Hoover report their correlations based on raw votes — not on shares of valid votes. Of course, even within a single precinct, MAS and MTS will both see larger numbers of votes at larger polling stations than at smaller polling stations that report earliest.

Figure 29
Correlation Between Matched Parties Varies By Support for Referendum



Sources and Notes: OEP (2016b), OEP (2016c), OEP (2017), OEP (2019b), OEP (2019c), OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations.

In **Figure 29**, we have divided the precincts into five groups of roughly equal size (919–949 precincts) based on opposition for the referendum in 2016 as a share of valid votes (a proxy for how "right" the precinct leans), so the groups are shown "left" to "right" politically. In more left-leaning precincts, the MAS and MTS shares substitute (negative correlation), and in more right-leaning precincts, MAS shares complement (positive correlation). Likewise, CC and 21F

correlate negatively in intensely opposition areas, but positively in areas supportive of the referendum.

Escobari and Hoover's basis for the parallel trends assumption is again very much in doubt.

Regardless, Escobari and Hoover expressly attribute any difference in trends between MAS-CC and MTS-21F as fraud. As MTS and 21F both have low vote shares throughout, their differences are also small and cannot trend much. Any meaningful trend in net votes for MAS will appear to be fraud.

Thinking in terms of the vote margins, changes in the minor-party margin (MTS - 21F) greatly limit how much the major-party margins may change without creating a "triple" difference. And the party shares are too small for the minor-party margins to change much. If MTS and 21F together claim only 5–7 percent of the valid vote, then their difference in shares can likewise be no more than 5–7 percentage points. If MTS got a zero share early and 21F zero share late, this would mean at most a 10–13 percentage point trend. In practice, the minor-party margins trended much less.

Table 6

Minor Parties, Being Small, Cannot Trend Much

	Major Margin (MAS-CC) (A)	Minor Share (MTS+21F) (B)	Minor Margin (MTS-21F) (C)	Difference (D=A-2B)
Early Polling Stations				
From Early Precincts	9.47	5.29	-2.76	-1.11
From Split Precincts	6.82	5.70	-3.28	-4.58
Late Polling Stations				
From Split Precincts	19.54	5.00	-2.57	9.54
From Late Precincts	52.12	6.73	-2.49	38.67
Total	10.56	5.49	-2.99	-0.42

Sources and Notes: OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations.

Thus, we have seen that the triple differences are not actually triple differences, and Escobari and Hoover lack justification in assuming that the parties show parallel trends. They fail to fully account for nickels before dimes. Furthermore, the resulting "triple" differences suffer the same Panic Pete problem as the difference-in-difference results: even assuming that fraud drives the lack of parallel trends, the approach cannot suggest which party the presumed fraud hurts and which it helps.

In short, Escobari and Hoover do not offer any convincing evidence based on the 2016 referendum data that there was fraud, let alone fraud sufficient to alter the official outcome of a first-round Morales victory.

Conclusions

In the end, Escobari and Hoover offer a series of models, none of which suffice to capture benign shifts in support across geography or elections. Rather than seeking explanation for such shifts, Escobari and Hoover simply label each failure of their models as evidence of fraud.

Escobari and Hoover claim to have estimated fraud of "2.50% of valid votes," which is more than four times the amount required to change the outcome of the election. They specifically claim to find 0.58 percentage points of fraud in the polling stations that reported after the interruption; alone, this would suffice to change the outcome. But as our Table 4 demonstrates, such extensive fraud requires implausible assumptions if the pre-interruption results are to be believed. To reach their conclusion, Escobari and Hoover must disbelieve the pre-interruption results. They justify this disbelief with a "triple-difference" model, which is not plausibly grounded.

- The "triple" in the triple difference comes from contrasting 2019 to 2016 to the minor-party results. But the 2016 referendum had no minor parties, so this cannot be a true triple difference, and Escobari and Hoover refuse to share data or methods that indicate otherwise. In fact, the 2016 data is completely irrelevant to the calculation of the triple difference. Replacing the actual 2016 results with random or otherwise utterly fabricated results has zero impact on the estimate.
- Regardless, the assumption of parallel trends between the major and minor parties is implausible on its face: minor-party support does not and largely cannot move percentage point for percentage point in step with major-party support.

Along with this "triple-difference" model, Escobari and Hoover offer a "difference-in-difference" model with nonparallel trends. This does not find meaningful "fraud" in the post-interruption results. Yet Escobari and Hoover offer no convincing argument as to why the trends should be parallel, why nonparallel trends should indicate fraud to the exclusion of all alternative explanations, or why such nonparallel trends should indicate fraud favoring Morales, specifically.

Surrounding this obviously flawed conclusion of widespread fraud, Escobari and Hoover present models all based on assumed parallel trends. Thus, Escobari and Hoover argued both that —

given the known shift in vote shares — there is fraud among late arrivals if the trends are parallel, and there is fraud everywhere if they are not. Whether directly or indirectly, they simply measure the difference in trends and label that "fraud."

Escobari and Hoover fail to meet any serious standard of evidence with respect to widespread fraud, let alone offer statistical evidence that the official result of Bolivia's 2019 presidential election should be questioned. In this, they join the OAS in destroying confidence in the electoral process.

References

Agence France-Presse (AFP). 2019. "Lo que debe saber para entender la crisis política en Bolivia que derive en la renuncia de Evo Morales." November 10. https://www.eluniverso.com/noticias/2019/11/10/nota/7598501/que-debo-saber-entender-crisis-politica-bolivia-derivo-renuncia-evo/

Agencia Boliviana de Información. 2019. "Pumari desmiente fraude en Potosí y aclara que lo que halló en un domicilio fue material usado en comicios." Archive Today, October 21. https://archive.ph/Wj64v

Almagro, Luis (@Almagro_OEA2015). 2019. "Comunicado del Grupo de Auditores del Proceso Electoral en #Bolivia: https://bit.ly/2NzuCIP." Twitter, November 10, 4:05 a.m. https://twitter.com/Almagro_OEA2015/status/1193454650831376385

Agencia de Noticias Fides (@noticiasfides). 2017. "#Ahora #ANF Resultados parciales de la votación correspondiente al Consejo de la Magistratura a nivel nacional: El 31,29% votos válidos; 15,30% blancos y 53,41% nulos." Twitter, December 3, 9:52 p.m. https://twitter.com/noticiasfides/status/937514848623185920

Agencia de Noticias Fides (ANF). 2019. "Víctor Borda suplica que liberin a su hermano tomado como rehén y renuncia a su curul." November 10. https://www.noticiasfides.com/nacional/politica/victor-borda-suplica-que-liberen-a-su-hermano-tomado-como-rehen-y-renuncia-a-su-curul-402315

BBC News. 2019. "Bolivia Election: Anger Mounts over Result Confusion." October 22. https://www.bbc.com/news/world-latin-america-50134451

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/

Chumacero, Rómulo A. 2019. "El camaleón, el mutante y Houdini: Resultados de las elecciones en Bolivia." October 26. http://rchumace.econ.uchile.cl/papers/cmh.pdf

Cuiza, Paulo. 2019. "Por fraude electoral detienen a la presidenta del TSE Choque y al exvocal Costas." La Razón, November 10. https://web.archive.org/web/20191111095015/http://www.la-razon.com/nacional/vocales-electoral-nacional-aprehendidos-electorales_0_3255274488.html

Collyns, Dan. 2019. "Bolivian president Evo Morales resigns after election result dispute." The Guardian, November 11. https://www.theguardian.com/world/2019/nov/10/bolivian-president-evo-morales-resigns-after-election-result-dispute

Comité Cívico Potosinista. 2019. Untitled post. Facebook, October 21, 2019. https://www.facebook.com/569728193073714/videos/410319093212525

Correo Del Sur. 2021. "Fiscalía exime de culpa a exvocales del TED Santa Cruz por caso Fraude." January 28. https://correodelsur.com/politica/20210128_fiscalia-exime-de-culpa-a-exvocales-del-ted-santa-cruz-por-caso-fraude.html

Corzo, Rodrigo. 2020. "21 Días: Cronología de los sucesos poselectorales de 2019 en Bolivia." Red de Líderes por la Democracia y Desarrollo (RELIDD). https://www.academia.edu/44943287/21_d%C3%ADas_Cronolog%C3%ADa_de_los_sucesos_poselectorales_de_2019_en_Bolivia

Deutsche Welle. 2019. "Bolivia President Evo Morales resigns." November 10. https://www.dw.com/en/bolivia-president-evo-morales-resigns/a-51195195

District of Columbia Board of Elections. 2020. "General Election 2020 - Certified Results." December 2. https://electionresults.dcboe.org/election_results/2020-General-Election

European Union's Election Expert Mission (EEM). 2019. "Bolivia 2019: Final Report, General Elections, 20 October 2019." https://ec.europa.eu/info/strategy/relations-non-eu-countries/types-relations-and-partnerships/election-observation/mission-recommendations-repository/missions/341

El Potosí. 2019. "Prenden fuego en la casa de Víctor Borda." November 10. https://elpotosi.net/local/20191110_prenden-fuego-en-la-casa-de-victor-borda.html

El Potosí Video. 2019. "Marco Pumari habla de la vigilia." YouTube video, October 21, 2:06. https://www.youtube.com/watch?v=fzVOI5xPBWY&t=0m37s

El Universal. 2019. "Evo Morales denuncia que incendiaron la casa de su hermana en Bolivia." November 9. https://www.eluniversal.com.mx/mundo/evo-morales-denuncia-que-incendiaron-la-casa-de-su-hermana-en-bolivia

Escobari, Diego and Gary A. Hoover. 2019. "Evo Morales and Electoral Fraud in Bolivia: A Natural Experiment Estimate." November 25. https://cepr.net/images/stories/reports/escobari_and_hoover_original.pdf

_____. 2020. "Evo Morales and Electoral Fraud in Bolivia: A Natural Experiment and Discontinuity Evidence." October 15. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3492928

EFE. 2019. "Renuncia el ministro de Minería, la primera dimisión en el Gobierno de Bolivia." November 10. https://www.eldiario.es/politica/renuncia-ministro-mineria-gobierno-bolivia_1_1264798.html

Erbol. 2022. "Exgobernado Víctor Hugo Vásquez vuelve al gobierno como Director de Soberanía Alimentaria." January 11. https://erbol.com.bo/nacional/exgobernador-v%C3%ADctor-hugo-v%C3%A1squez-vuelve-al-gobierno-como-director-de-soberan%C3%ADa-alimentaria

Ethical Hacking. 2019. "Informe Consolidado de Producto 1 y Producto 2 Línea de Tiempo de la Consultoría al OEP." https://cepr.net/wp-content/uploads/2020/03/EHC-REP-Consolidado-Timeline-OEP-No-06-v1.4.pdf

Ministerio Público, Fiscalía General del Estado. 2021. "Resolución de Sobreseimiento No. 06/2021." July 27. https://drive.google.com/file/d/1qmgB3Q3ZpXtVvLhQJj-rJHoL7MuIXQML/view

Gelman, Andrew. 2019. "Should We Be Suspicious of the Vote Counting in Bolivia?" Statistical Modeling, Causal Inference, and Social Science, November 25. https://statmodeling.stat.columbia.edu/2019/11/25/should-we-be-suspicious-of-the-vote-counting-in-bolivia/

Gobierno de México. 2019. "Position of the Government of Mexico on Granting Asylum to Evo Morales." November 11. https://www.gob.mx/sre/prensa/position-of-the-government-of-mexico-on-granting-asylum-to-evo-morales

Human Rights Watch. 2020. September 11. "Justice as a Weapon: Political Persecution in Bolivia." https://www.hrw.org/report/2020/09/11/justice-weapon/political-persecution-bolivia

Idrobo, Nicolás, Dorothy Kronick, and Francisco Rodríguez. 2022. "Do Shifts in Late-Counted Votes Signal Fraud? Evidence from Bolivia." *Journal of Politics* 84.4. October. https://doi.org/10.1086/719639

Inocente, Mauricio (*@Mojris*). 2019. "Queman Tribunal electoral en #Sucre #FraudeEnBolivia #21f #EleccionesGenerales2019 #EleccionesBolivia2019 #Elecciones2019 #EleccionesBo2019 #bolivia." Twitter, October 21, 9:15 p.m. https://twitter.com/Mojris/status/1186450855400235010

La Razón Digital (@LaRazon_Bolivia). 2016. "#Referendo2016 #21F Datos Preliminares transmisión rápida de actas. Al 83,3%, el Sí obtiene 46,2% y el No 53,8%." Twitter, February 22, 8:23 p.m. https://twitter.com/LaRazon_Bolivia/status/701940460328189953

La República. 2019. "Bolivia: Incendian la casa de la hermana de Evo Morales [VIDEO]." November 10. https://larepublica.pe/mundo/2019/11/10/protestas-en-bolivia-en-vivo-ultimas-noticias-casa-de-hermana-evo-morales-esther-incendio-gobernador-victor-vasquez-oruro-video-mdga/

Long, Guillaume, and Lola Allen. 2019. "Evo Morales's Life was in Danger, and He Almost Didn't Make it Out of Bolivia." Washington, DC: Center for Economic and Policy Research, November 20. https://cepr.net/evo-s-odyssey-how-mexico-got-morales-out-of-bolivia/

Los Tiempos. 2019a. "TSE dice que se paralizó el TREP para 'evitar la confusión' con resultados de departamentos." October 20. https://www.lostiempos.com/actualidad/pais/20191020/tse-dice-que-se-paralizo-trep-evitar-confusion-resultados-departamentos

_____. 2019b. "Un triunfo de la esperanza." October 21. https://web.archive.org/web/20191028185324/https://www.lostiempos.com/actualidad/opinion/20191021/editorial/triunfo-esperanza

Louisiana Department of State. 2019. "Official Results." November 16. https://voterportal.sos.la.gov/ElectionResults/ElectionResults/Data?blob=20191116/csv/ByParish_57627.csv

Mancilla, Alfredo Serrano. 2021. "Evo: Operación Rescate: Una trama geopolítica en 365 días." Argentina: Penguin Random House Grupo Editorial Argentina.

Mayora, Fernando and Benjamín Rodríguez. 2016. "Urnas y democracia directa: Balance del Referendo Constitucional 2016." Órgano Electoral Plurinacional (OEP) Bolivia, June. https://pioresnada.files.wordpress.com/2016/12/urnas-y-democracia_directa-referendo-2016.pdf

Mesa Gisbert, Carlos D (@carlosdmesag). 2019. "¡Gracias Bolivia! ¡Vamos a salvar la democracia!" Twitter, October 20, 8:57 p.m. https://twitter.com/carlosdmesag/status/1186084073938341889

NEOTEC. 2019. "Informe Final Elecciones Generales 2019." October 28. https://cepr.net/wp-content/uploads/2020/03/Informe-de-Neotec-sobre-las-elecciones-en-Bolivia.pdf

New York Times, The. 2020. "North Carolina Presidential Election Results." November. https://www.nytimes.com/interactive/2020/11/03/us/elections/results-north-carolina-president.html

North Carolina State Board of Elections (@NCSBE). 2020. "Results will be released when the final poll closes at 8:15 p.m." Twitter, November 3, 8:03 p.m. https://twitter.com/NCSBE/status/1323792958047588352

_____. 2022. "Election Night Reporting Timeline." https://www.ncsbe.gov/results-data/election-results/election-night-reporting-timeline

OpenDemocracy. 2019. "The bible makes a comeback in Bolivia with Jeanine Añez." November 20. https://www.opendemocracy.net/en/democraciaabierta/quién-es-jeanine-añez-y-por-qué-desprecia-los-pueblos-indígenas-de-bolivia-en/

Opinión. 2016. "Al 83.1% de las actas escrutadas el No tiene 53.8% y el Sí 46.2%." February 21. https://www.opinion.com.bo/articulo/referendo/831-actas-escrutadas-tiene-538-s-iacute-462/20160221200100543821.html

Organization of American States (OAS). 2019a. "Electoral Integrity Analysis: General Elections in the Plurinational State of Bolivia: October 20, 2019 Final Report." https://www.oas.org/fpdb/press/Audit-Report-EN-vFINAL.pdf

_____. 2019b. "Organization of American States: Electoral Observation Missions: General Elections 2019 and 2020: Plurinational State of Bolivia Final Report: General Elections October 20, 2019 Final Report." ACE Electoral Knowledge Network. https://aceproject.org/ero-en/regions/americas/BO/bolivia-general-elections-final-report-2019

_____. 2019c. "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

_____ (@OEA_oficial). 2019d. "La Misión de Observación Electoral de la Organización de los Estados Americanos (OEA) en #Bolivia llama a que se respete la voluntad de la ciudadanía. Declaración: https://bit.ly/2ByWN3Q #OEAenBolivia." Twitter, October 21, 9:31 p.m. https://twitter.com/OEA_oficial/status/1186455002337959937

_____. 2019e. Statement of the Group of Auditors Electoral Process in Bolivia." November 10. https://www.oas.org/en/media_center/press_release.asp?sCodigo=E-099/19

_____ (@OAS_official). 2019f. "A coup d'état in #Bolivia? Yes. When they tried to steal the elections of October 20 and perpetuate @evoespueblo in power. Part 1 of the speech of Secretary General @Almagro_OEA2015 to the #OAS Permanent Council. #OEAenBolivia." Twitter, November 12, 6:00 p.m. https://twitter.com/oas_official/status/1194389549037830145

Órgano Electoral Plurinacional (OEP) Bolivia. 2016a. "Resultados Oficiales: Referendo Constitucional 2016." March 13. https://www.oep.org.bo/wp-content/uploads/2017/02/resultados_referendo_constitucional_2016.pdf

_____. 2016b. "Votos_totales." https://atlaselectoral.oep.org.bo/descarga/38/votos_totales.csv

_____. 2016c. "Votos_totales." https://atlaselectoral.oep.org.bo/descarga/83/votos_totales.csv

_____. 2017. "Referendum 2016: General Estadistico." https://www.oep.org.bo/wp-content/uploads/2017/02/padron_ref_const_2016.xlsx

_____. 2019a. "Más de 8.000 notarias y notarios darán fe de los actos electorales en las Elecciones Generales 2019." Fuente Directa, September 30. https://fuentedirecta.oep.org.bo/noticia/mas-de-8-000-notarias-y-notarios-daran-fe-de-los-actos-electorales-en-las-elecciones-generales-2019/

_____. 2019b. "Elecciones Generales 2019: Habilitados, Inhabilitados y Deputados por Recinto." https://www.oep.org.bo/wp-content/uploads/2019/09/Estadisticas_Exterior_EG_2019.xlsx

_____. 2019c. "Elecciones Generales 2019: Habilitados por mesa por grupo Etario." https://www.oep.org.bo/wp-content/uploads/2019/09/Estadisticas_Nacional_EG_2019.xlsx

_____. 2020a. "Bolivia (2019)." April 7. Internet Archive, https://archive.org/details/election_bolivia_2019

_____. 2020b. "Reports_computo.zip." April 7. Internet Archive, https://archive.org/download/election_bolivia_2019/reports_computo.zip/acta.2019.10.25.21.09.30.xlsx

_____. 2020c. "Reports_trep.zip." April 7. Internet Archive, https://archive.org/download/election_bolivia_2019/reports_trep.zip/acta.2019.10.20.19.40.57.xlsx

_____. 2022. "Cómputo Electoral Elecciones Generales 2019." August 6. https://computo2019.oep.org.bo/simobol_2019_3_Computo/PubResul/acta.2022.03.28.15.25.09.xlsx

Página Siete. 2019a. "Al 100% de conteo rápido, Viaciencia ratifica segunda vuelta entre Evo y Mesa." October 20. https://web.archive.org/web/20201112021415/https://www.paginasiete.bo/nacional/2019/10/20/al-100-de-conteo-rapido-viaciencia-ratifica-segunda-vuelta-entre-evo-mesa-235009.html

_____. 2019b. "Murillo a Quintana y Raúl García Linera: Que empiecen a corer porque los vamos a agarrar." https://web.archive.org/web/20191114063354/https://www.paginasiete.bo/seguridad/2019/11/13/murillo-quintana-raul-garcia-linera-que-empiecen-correr-porque-los-vamos-agarrar-237335.html

_____. 2020a. "Nuevo estudio muestra fraude antes de la caída del TREP." October 15. https://www.paginasiete.bo/nacional/nuevo-estudio-muestra-fraude-antes-de-la-caida-del-trep-FGPS271576

_____. 2020b. "TSE determina la suspensión del Direpre." October 17. https://www.paginasiete.bo/nacional/tse-determina-la-suspension-del-direpre-BGPS271891

Red Uno de Bolivia. 2019. "Sucede ahora mismo en el Tribunal Electoral Departamental de Sucre." Facebook, October 21. https://www.facebook.com/watch/?v=2578874505535164

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/

_____. 2020a. "Bolivia-Elections-2019." GitHub, June 8. https://github.com/ViscidKonrad/Bolivia-Elections-2019/commits/master/datos_trep_computo_eg2019.rar

_____. 2020b. "The OAS Misused Its Own Data to Help Fabricate Its Accusation of Fraud Against Evo Morales." Washington, DC: Center for Economic and Policy Research, September 8. https://cepr.net/the-oas-misused-its-own-data-to-help-fabricate-its-accusation-of-fraud-against-evo-morales/

_____. 2020c. "What Criticisms of Bolivia's 2019 Elections Continue to Get Wrong." Washington, DC: Center for Economic and Policy Research, December 9. https://cepr.net/report/what-criticisms-of-bolivias-2019-elections-continue-to-get-wrong/

_____. N.d. "Bolivia-Escobari-Hoover." Github. https://github.com/ViscidKonrad/Bolivia-Escobari-Hoover

RTP Bolivia. 2020. "Ministerio Público: 'Paralización del Sistema TREP no fue planificada'." October 7. http://www.rtpbolivia.com.bo/2020/10/07/ministerio-publico-paralizacion-del-sistema-trep-no-fue-planificada/

Silva (@bolsoringo22). 2022. Untitled post. Twitter, May 11, 2022, 2:22 p.m. https://twitter.com/bolsoringo22/status/1524454923647258624

Tribunal Supremo Electoral (TSE). 2019a. "Reglamento para las Elecciones Generales 2019." La Paz: Órgano Electoral Plurinacional. https://www.oep.org.bo/wp-content/uploads/2019/05/Reglamento_Elecciones_Generales_2019.pdf

_____. 2019b. "Resultados Preliminares de Las Elecciones Nacionales (TREP)." Youtube video, October 20, 2:54. https://youtu.be/FV0GEzRKkOY

Universidad de Salamanca. 2021. "Análisis de Integridad Electoral: Elecciones Generales en el Estado Plurinacional de Bolivia 20 de Octubre de 2019." https://cepr.net/wp-content/uploads/2022/09/1_5170467409735188784.pdf

Valverde, Rafael Lopez. 2020. "Carlos Mesa se declara en segunda vuelta." YouTube video, February 24, 7:49. https://youtu.be/pkKb4cvYYiI

Villamontes, Miguel Angel Roca. 2019. "Tras el informe de la OEA, Evo anuncia nuevas elecciones en Bolivia." El Deber, November 10. https://eldeber.com.bo/pais/tras-el-informe-de-la-oea-evo-anuncia-nuevas-elecciones-en-bolivia_156154

Villanueva R, Amaru (@amaruvr). 2019a. "Lo anterior nos permite hacer una estimación del voto final en base al TREP y el dato que reporta Carlos Mesa. Bajo este escenario, el MAS obtendría una diferencia del 10,9% sobre CC. Si el voto por el MAS en el área rural fuera aún mayor (65%), la diferencia sería 47,9% vs 35,9%." Internet Archive, October 21, 1:50 p.m. https://web.archive.org/web/20191021210238/https://twitter.com/amaruvr/status/1186384222321332225

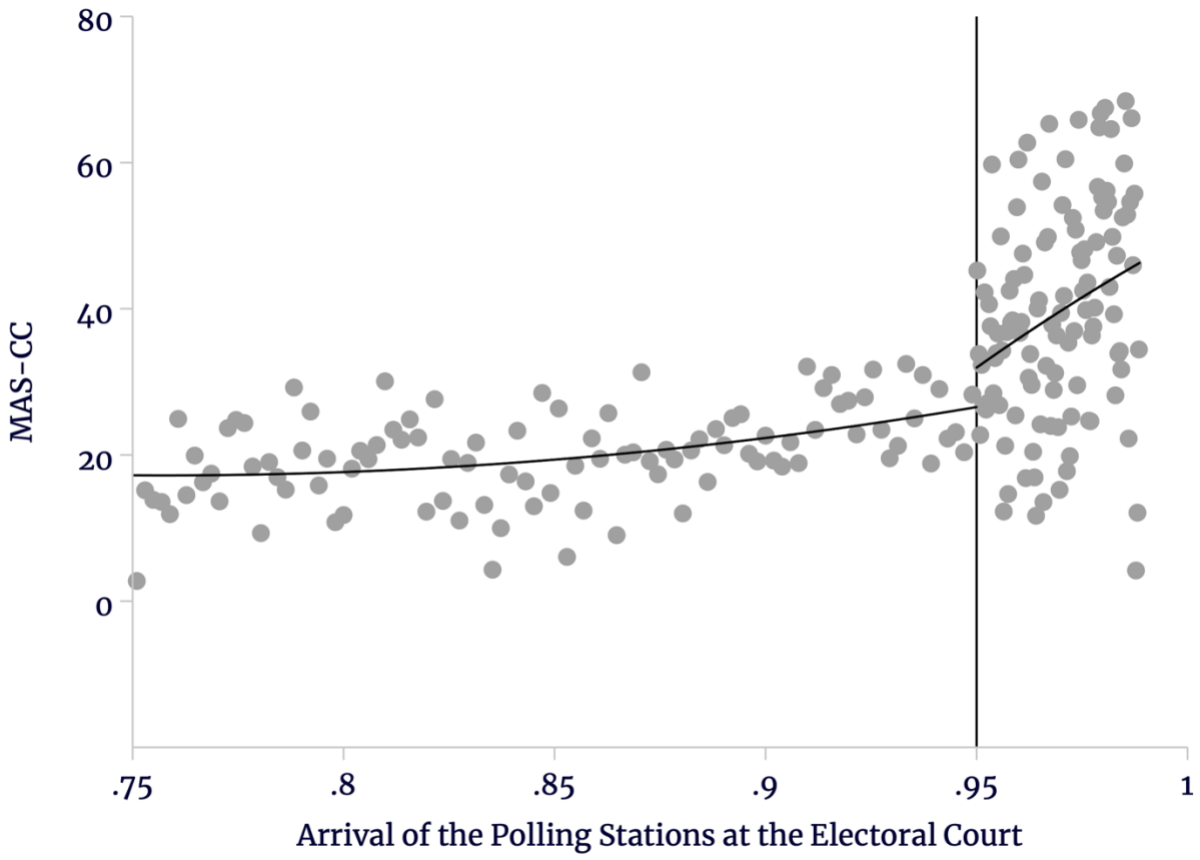
_____. 2019b. "Entiendo que hay que seleccionar 'mundo' porque el default es Bolivia, y sale unos decimales menos." Internet Archive, October 21, 2:12 p.m. https://web.archive.org/web/20191021212254/https:/twitter.com/amaruvr/status/1186389800171032578

Appendix A: Regression Discontinuity Analysis

Escobari and Hoover conclude their series of analyses with a regression discontinuity analysis. In many ways, this is a step backward in terms of understanding the election. In effect, it is a return to their earliest difference estimates, as there are no considerations given whatsoever to geography, nor are there considerations of the 2016 results. Simply, rather than assuming a constant difference between MAS and CC vote shares among the early arrivals, Escobari and Hoover allow a trend. They assert that absent fraud, the results of early arrivals should smoothly transition into the results of later arrivals.

However, there are obvious problems with this approach. In **Figure A1**, we see a replication of their Figure 5, illustrating an apparent break in trend for the last 5 percent of the vote.

Figure A1
Reproduction of Escobari and Hoover's Figure 5



Sources and Notes: Rosnick (2020a), TSE (2019a), and authors' calculations.

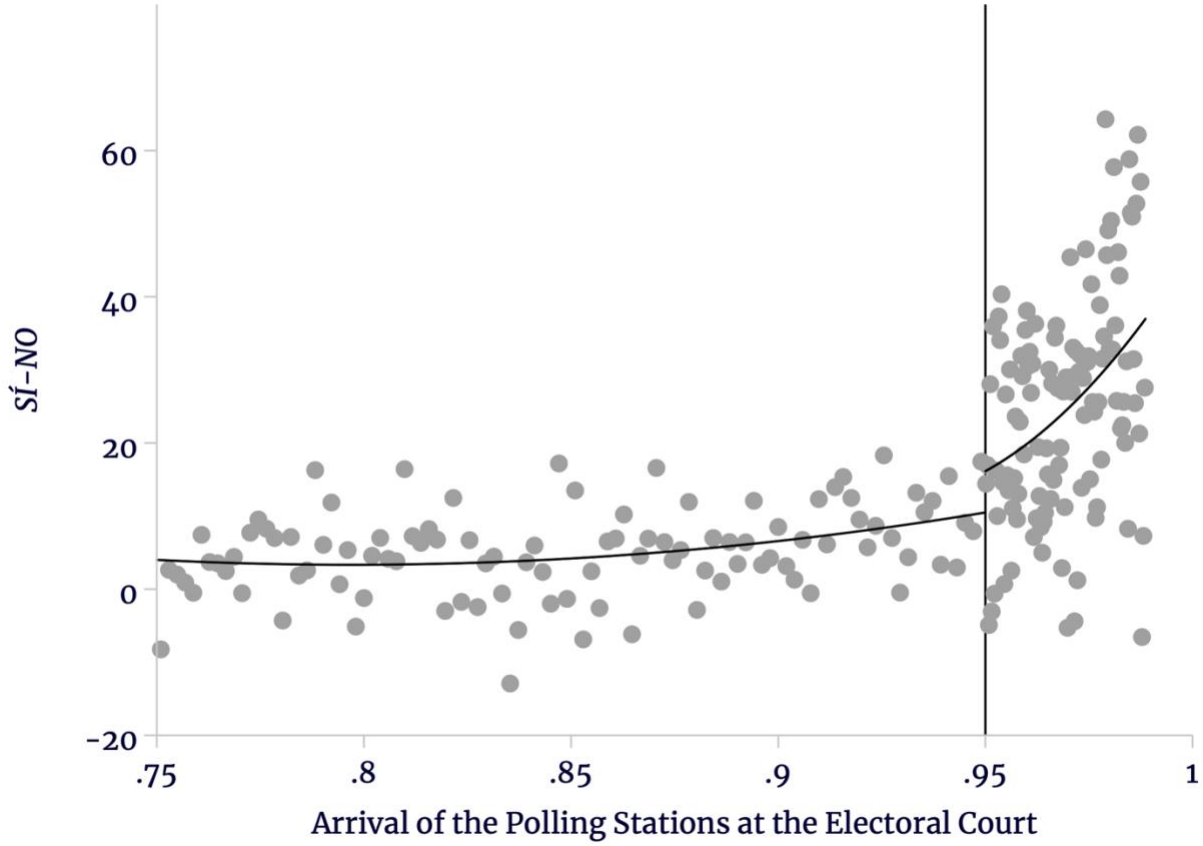
There is no clear explanation for why the focus has shifted to the last 5 percent of the count, except to say: "Our focus on the 0.95 cutoff follows the discussion of the statistical section of the (Organization of American States, 2019, p. 88) report that shows graphical evidence of an apparent discontinuity at 0.95."[126]

As we have pointed out previously,[127] Nooruddin, on behalf of the OAS, also offers no explanation for this choice. Furthermore, the 95 percent mark does not appear to be interesting, relative to other choices of break. For example, the same approach shows Mesa's share of the vote jumping sharply at 80 percent. Nooruddin's findings are dependent on the use of unequal bandwidths on either side of the break — as computed, the "left" side is not actually comparable to the "right." Finally, even if a break were to be convincingly argued, the results of the last 5 percent of the count are predictable. If we didn't know the actual results to the right of the break, but substituted instead our predictions for those vote shares and applied the same method, we find the determination of a "discontinuity" is expected — not unexpected. In any case, the OAS erroneously suggests a break in the last 5 percent of actas verified, not the last 5 percent that were transmitted.

However, there is an even simpler piece of evidence suggesting that the analysis does not demonstrate fraud. In **Figure A2**, we perform the same analysis as in Figure A1; the only difference is that we employ 2016 data. The similarity between the two would be striking were it not already abundantly established that what we are seeing is caused by the ordering of the count, i.e., by construction common to both. *Nickels before dimes*.

126 Escobari and Hoover (2020), 26.
127 Rosnick (2020c).

Figure A2
Reanalysis Based on 2016 Referendum Results

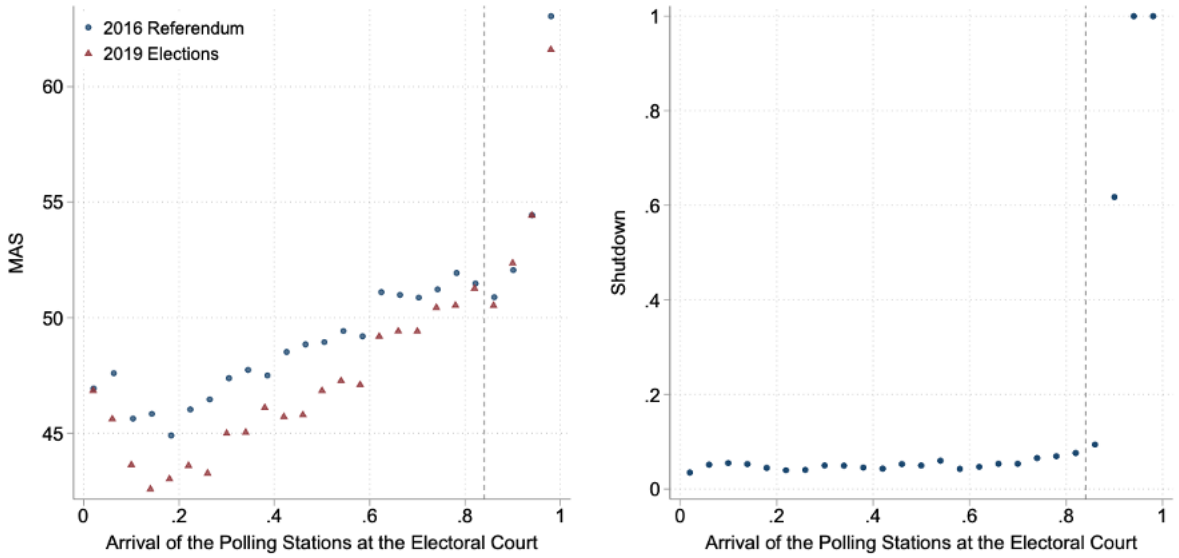


Sources and Notes: OEP (2016b), OEP (2016c), OEP (2017), OEP (2019b), OEP (2019c), OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations.

Appendix B: Escobari and Hoover's Figure 4

Figure B
Escobari and Hoover's Figure 4

Figure 4: MAS vote share upon arrival at the Electoral Court



Notes: The left-hand side shows MAS (or YES) votes as a function of the arrival of the polling stations at the Electoral Court, (ARRIVAL). The right-hand side shows the share of shutdown polling stations as a function of ARRIVAL.

Sources and Notes: Escobari and Hoover (2020).

Appendix C: Handling Within-Precinct Trends

In the synthetic data, we are simply assuming that within-precinct trends indicate fraud. We are then left with a Panic Pete problem: where along the arrival is the fraud located? To illustrate, we have Figure 13.

We may try to isolate the within-precinct trend through *stratification*. We may examine each precinct in isolation, estimate the trend for each precinct, and average the results to see if there is an overall tendency. Unsurprisingly, this is a major approach of Idrobo, et al. They examine Morales's margin (Morales's vote share minus that of runner-up Carlos Mesa) across the share of the vote publicly reported at any time (as opposed to arrival — the time a precinct successfully reports its results to the election offices). They find an unexplained within-precinct trend of approximately one-half of one percentage point additional (net) support for Morales from the start to finish — or 0.005 percentage points per percentage point of progress.[128] Though interesting, this is unimportant in political terms.

Another approach would be to divide the precincts by support for the referendum. This has the added advantage that in the actual data, the individual trends for precincts with few polling stations are highly sensitive to few votes or to very small changes in timing.[129] In precincts with no polling stations, the trend is impossible to compute. In stratifying by the results of another election (assumed free of fraud) rather than individual precincts, we effectively pool multiple, but similar, precincts. This allows precincts with single polling stations to contribute to the analysis. The disadvantage is that by combining precincts in each stratum, we fail to capture all cross-precinct differences that may correlate with arrival. Cross-precinct, as well as within-precinct trends, within strata may remain.

For example, on average relative to the referendum, Morales lost support in capital city precincts, and much less in rural precincts. Thus, we might expect that within a narrow range of support for the referendum, support for Morales was higher in rural precincts than in urban. As the rural precincts tended to report later, this would create within-stratum trends over the count.

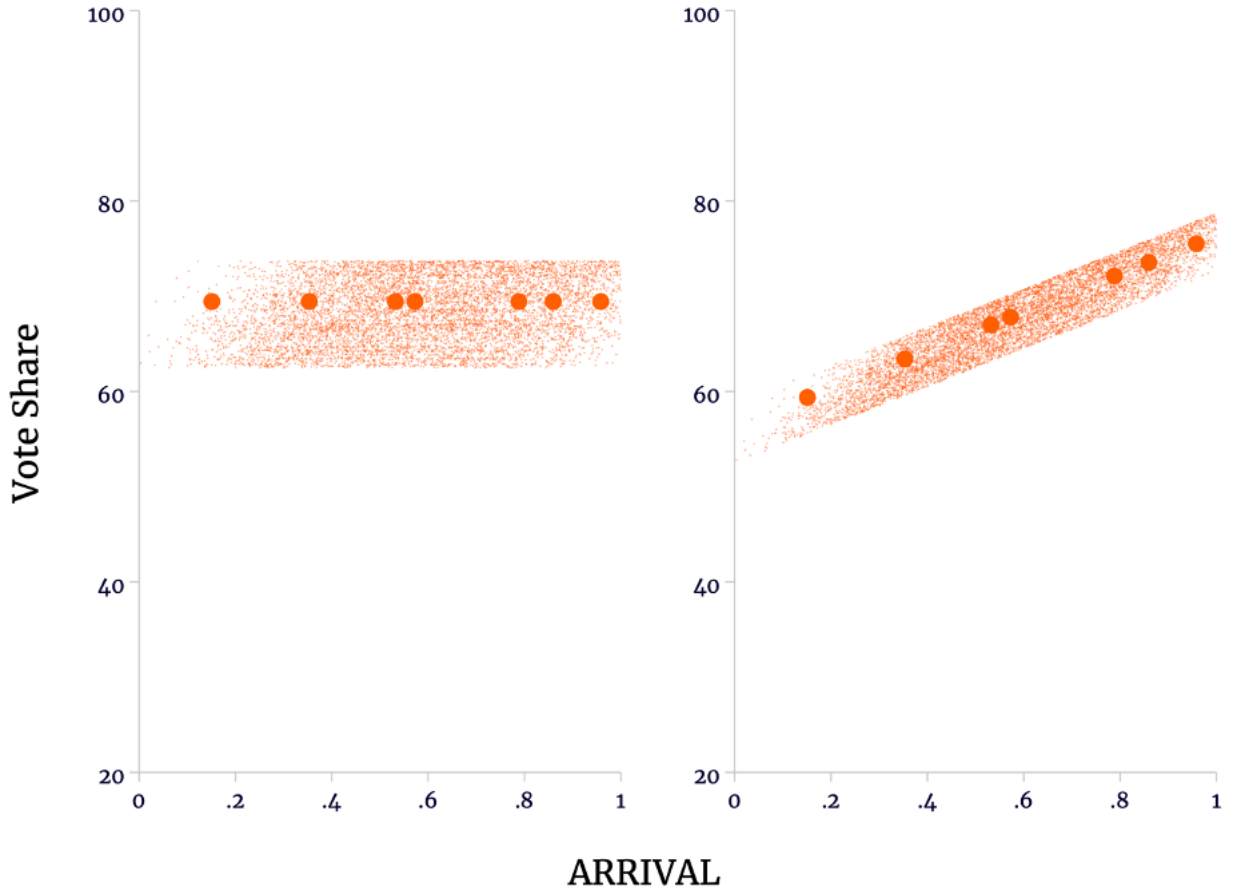
128 Idrobo, Kronick, and Rodríguez (2022), Figure 4b.

129 Suppose a precinct with only two polling stations transmits both results at the same time so that the arrival variable differs by only 1/35,000. If the first polling station reported 4 of 10 votes for one candidate, and the second reported 100 of 200, then the trend is (50% - 40%)/(1/35000) = 35000/10, or 3,500 percentage points of support per percentage point of progress in transmission. If just one more person in the first station had voted for the candidate, there would be zero trend.

Stratification on referendum results alone does not suffice. However, the referendum results do explain much of the overall trend.

In **Figure C**, we illustrate by restricting the synthetic data of Figure 5 to precincts within a narrow stratum of support for the referendum. The election vote shares are, of course, unrestricted. Because the division limits the amount of cross-precinct variation and there is no additional source of variation, the within-precinct trend dominates the trend within the stratum.

Figure C
Illustration of Stratification Exposing Within-Precinct Trend



Sources and Notes: Authors' calculations.

For purposes of the synthetic data we produce, it is sufficient to recognize (*a*) that we may generate data with trends within or merely across precincts, or both, or neither; and (*b*) that care must be taken to ensure that these different trends are identified separately.

Appendix D: Escobari and Hoover's Table 2

Table D

Replication of Escobari and Hoover's Table 2

	CC	MAS	MAS-CC			
	(A)	(B)	(C)	(D)	(E)	(F)
Variable						
SHUTDOWN	-8.268	7.975	16.26	7.243	6.762	0.377
	(0.324)	(0.343)	(0.653)	(0.437)	(0.464)	(0.194)
Constant	36.86	46.69	9.830	11.28	11.36	12.39
	(0.136)	(0.134)	(0.266)	(0.162)	(0.151)	(0.063)
Fixed Effects†						
Municipality				129.6		
Locality					23.49	
Precinct						124.7
Observations	34,529	34,529	34,529	34,529	34,529	34,529
R-squared	0.017	0.016	0.017	0.640	0.740	0.958

Sources and Notes: OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a), and authors' calculations. Dependent variables are percentages of *Válidos En Acta* (frequently missing or otherwise misreported on the tally sheets) and not of official valid votes. Standard errors in parenthesis are robust. Differences from Escobari and Hoover noted in red.
† F-test for fixed effects computed based on non-robust errors.

Appendix E: Escobari and Hoover's False and Misleading Claims About Argentina

- Claim 1: "137 ballot boxes recorded over 90% of their votes for Morales."[130]

This is false. There were only 113 such ballot boxes out of 711 total. If that seems high, consider that 10 of the 66 polling stations in the United States and 4 of 29 in Italy saw Morales getting less than 10 percent of the vote.

- Claim 2: "while participation rates increase on average 4.8% between 2016 and 2019 across all ballot boxes, for Argentina our data shows they increased by 154.6%."[131]

In fact, the rate of turnout from abroad in 2016 (31.3 percent) was well below domestic (86.7 percent), and turnout in Argentina was particularly low — only 25.6 percent of eligible voters cast a ballot. In 2019, the rate of turnout from abroad nearly doubled to 61.6 percent, but the 2016 referendum was exceptional — not the rule. Turnout in Argentina ran in the low 60s for general elections in 2014 and 2020, as well as in 2019. The OEP offered explanations for the particularly high rate of absenteeism, including nonmandatory participation, general disinterest, and seasonal travel.[132]

Given its particularly low base in 2016, the increase in turnout from Argentina is not large.

- Claim 3: "There is even a ballot box where Morales officially recorded 153% of the valid votes." (Added footnote: "The official Electoral Court records show seven polling stations where MAS votes exceed the total number of valid votes.")[133]

This is false. There is not a single polling station where this is true in the official count for any party. In fact, at every polling station, the official total number of valid votes is exactly equal to the sum of the official votes for each party. The polling station to which Escobari and Hoover refer is mesa number 1579, where MAS received 167 of 189 valid votes.

130 Escobari and Hoover (2020), 20.
131 Escobari and Hoover (2020), 20.
132 Mayora and Rodríguez (2016).
133 Escobari and Hoover (2020), 20.

The "153%" statistic comes from dividing 167 by 109 — the "Válidos En Acta." For nearly every polling station in the official data, the "Válidos En Acta" is an initial transcription of the "Votos Válidos" check on the acta — the summary record produced at the polling station. This number has no official standing, but rather is supposed to be used by the jurors to help ensure that the count is free of math errors. However, the figure is often incorrect due to arithmetic error or misunderstanding on the part of the jurors as to what number belongs there. Sometimes, it is missing from the acta altogether. For purposes of determining the election outcome, it is the sum of officially recognized valid votes for parties that is legal, not the number reported on the acta. In the case of mesa 1579, e.g., the middle digit "8" had been transcribed incorrectly as "0."

In fact, the sum of "Válidos En Acta" throughout the election is 6,137,778. By contrast, the official count of valid votes is 6,137,671. Escobari and Hoover agree to this latter sum, although they employ the former numbers in computing party vote shares and margins.[134] We have previously pointed out their inconsistent use of official and unofficial valid vote totals.

134 Escobari and Hoover (2020), Table 1.

Appendix F: More Considered Estimation of the Final Election Results

In Table 4 above, we produced one very simple forecast of the final election results based on (1) the official results at polling stations included in the TSE announcement, and (2) the total number of valid votes officially recorded at all polling stations. This forecast made additional use of the 2016 referendum results to estimate the 2019 results in precincts for which the TSE had reported no results whatsoever.

The total number of valid votes officially recorded at late polling stations would not be known with certainty at the time of the TSE announcement; these we had presumed correct in order to simplify the analysis.

In the following estimation, we trade simplicity for additional rigor. Here, we forecast a range of possible outcomes consistent with the data as was available at the time of the TSE announcement: the results as announced, and the number of voters eligible at each polling station. The number of valid votes in each outstanding precinct, in particular, are forecast based on the patterns of turnout reported at the time of the announcement. No use is made of any prior election results, nor are the forecasts based on any results reported after the announcement.

To accomplish this, we assume that voters are identical within a given precinct. The voters (actual or potential) are divided into four groups: for Morales, for Mesa, for any other candidate (other), and any nonvalid (blank, null, or nonvoting). Within any precinct with any polling stations included in the TSE announcement, the distribution of voters is assumed Dirichlet-multinomial with prior alpha = (1,1,1,1). For precincts with no results available at the time of the TSE announcement (a "target" precinct), the parameters of the Dirichlet-multinomial draw are imputed at random from among any representative precincts. In turn, representative precincts are determined as follows:

If a target precinct lies in a zone containing announced results, then the precincts of that zone with any announced results are the representative precincts for the target. Otherwise, if a target zone lies in a district containing announced results, then the zones of that district with any announced results are the representative zones for the target. The process of selecting representative geographies expands as necessary until representative subgeographies may be selected.

To illustrate, the precinct of Unidad Educ. Eduardo Lopez O E. Don Bosco had no polling stations included in the results announced by the TSE. Thus, we look for other precincts in Zona 45, Distrito 33, with available data. Finding none, we expand the search to all of Distrito 33 (Champa Rancho). We see that only one polling station located in Champa Rancho was included in the TSE announcement — that of U. E. Alcides Arguedas A. We therefore infer that Champa Rancho's *zona* San Juan Bosco must be the representative zone for O E. Don Bosco. Had results existed in an additional zone of Distrito 33, then we would select a representative zone at random between the two. Given that San Juan Bosco is the representative zone, the representative zone is (trivially) Alcides Arguedas A. Thus, results at O E. Don Bosco are imputed as if those polling stations were an extension of the precinct Alcides Arguedas A.

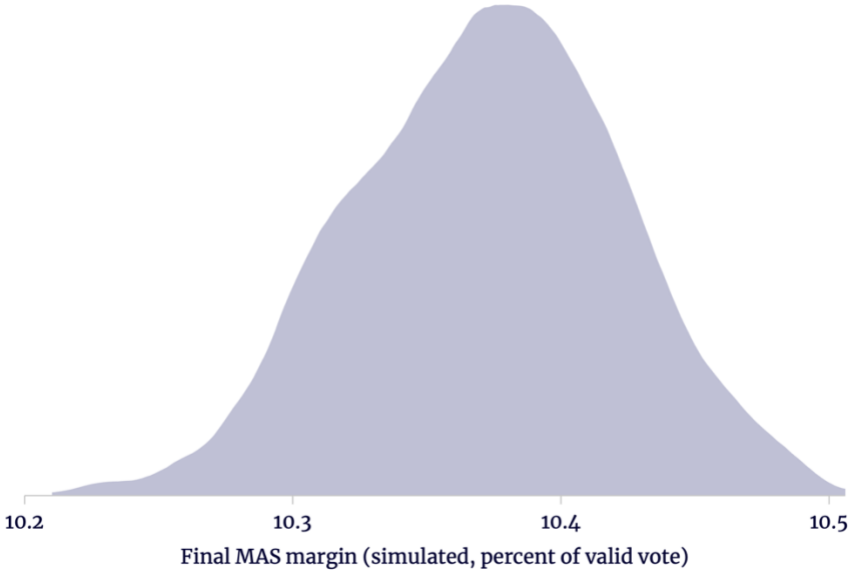
To take a more extreme example, no results from Colombia were reported in the TSE announcement. Thus, we choose at random a representative country for Colombia from among the many countries with early results. Suppose we choose the United States. This leaves us five states, plus the District of Columbia, from which to draw at random a representative department. If we choose Virginia, we have Arlington and Fairfax from which to draw a representative province. The only available municipality in Arlington is Arlington itself, the only available district in Arlington is Distrito Arlington, and the only available zone in that district is Zona Virginia. Thus, having selected Arlington at random as the representative province, we may only choose Drew Model Elementary School or Wakefield High School as the representative precinct for our target (the embassy in Bogotá, Colombia). We flip a coin (choosing, say, Wakefield) and perform a Dirichlet-multinomial draw for the embassy based on the parameters for the representative precinct (Wakefield High).

In this way, we impute voters at all polling stations not announced by the TSE.[135] Because the representatives and the voters themselves are selected at random, we may repeat the entire process many times and produce different results. This yields a range of credible projections for the final results. In **Figure F**, we see a distribution of final margins based on 1,000 such imputations.

135 One very technical note: where the number of available polling stations is very limited, we may — again at random — choose to select a representative from a wider geography. For example, when we determined that Alcides Arguedas A. was the only available precinct to represent O E. Don Bosco, because Alcides Arguedas A. is the only precinct in Distrito 33 with available data, this was an oversimplification. We may sometimes choose to go beyond Distrito 33 and follow the same procedure using all of the available districts within the locality of Cochabamba as possible representatives for O E. Don Bosco. This makes the results less certain, with the added benefit of avoiding overconfidence where information is limited. A fuller Bayesian analysis is possible, but costly. Adhering strictly to the data available within a geography does not importantly change the projections in this case, but may make results much less certain when data is less densely available.

Figure F

Projected Final Results Based on Monte Carlo Imputation



Sources and Notes: OEP (2020b), OEP (2020c), Rosnick (2020a), TSE (2019a) and authors' calculations. One thousand iterations based on results from 28,975 of 34,555 mesas.

We see that the full range of margins run from 10.21 to 10.51 percentage points of the valid vote, with 80 percent of the imputations falling between 10.30 and 10.44 percentage points. That is, the approach offers considerable evidence that in fact Morales was going to win in the first round based on the data available at the time of the TSE announcement.