On Unfounded Claims of Electoral Fraud in Bolivia

Nicolás Idrobo[†] Dorothy Kronick[‡] Francisco Rodríguez[§]

November 27, 2024

Abstract

An article recently published in this journal claims to present statistical evidence of fraud in Bolivia's controversial 2019 presidential election. These claims are significant not only for our understanding of a pivotal moment in Latin American politics but also because, as the authors note, their methods might inform how researchers investigate fraud in other cases. We explain why the evidence does not support the authors' conclusions. They claim to find evidence of fraud based on: (1) a difference-in-differences, (2) a simple difference, and (3) regression discontinuity. But (1) the pre-trends are converging in the difference-in-differences, (2) there are many benign explanations for the simple difference, and (3) the regression discontinuity uses an arbitrarily chosen cutoff at which placebo outcomes are not smooth. Our objective is both to correct the record about this specific election and, more generally, to reiterate the risks of ad hoc election forensics.

[†]Department of Political Science, University of Pennsylvania, idrobo@sas.upenn.edu

[‡]Corresponding author (kronick@berkeley.edu). Goldman School of Public Policy, University of California, Berkeley

[§]Josef Korbel School of International Studies, University of Denver, francisco.rodriguez4@du.edu

In a paper recently published in this journal, Escobari and Hoover (2024) claim to find statistical evidence of fraud in Bolivia's contested 2019 presidential election. These claims are significant not only for our understanding of a pivotal moment in Bolivian politics, Latin American politics, and international relations, but also because, as the authors note, their methods might inform how researchers investigate fraud in other cases.

In this response, we explain why Escobari and Hoover's evidence does not support their conclusions. They claim to find evidence of electoral fraud based on (1) a difference-in-differences, (2) a simple difference, and (3) regression discontinuity. But, as we establish below, (1) the pre-trends are converging in the difference-in-differences, (2) there are many benign explanations for the simple difference, and (3) the regression discontinuity uses an arbitrarily chosen cutoff at which placebo outcomes are not smooth.

In what follows, our objective is both to clarify the record about this specific election and, more generally, to comment on the substantive political risks of ad hoc election forensics.

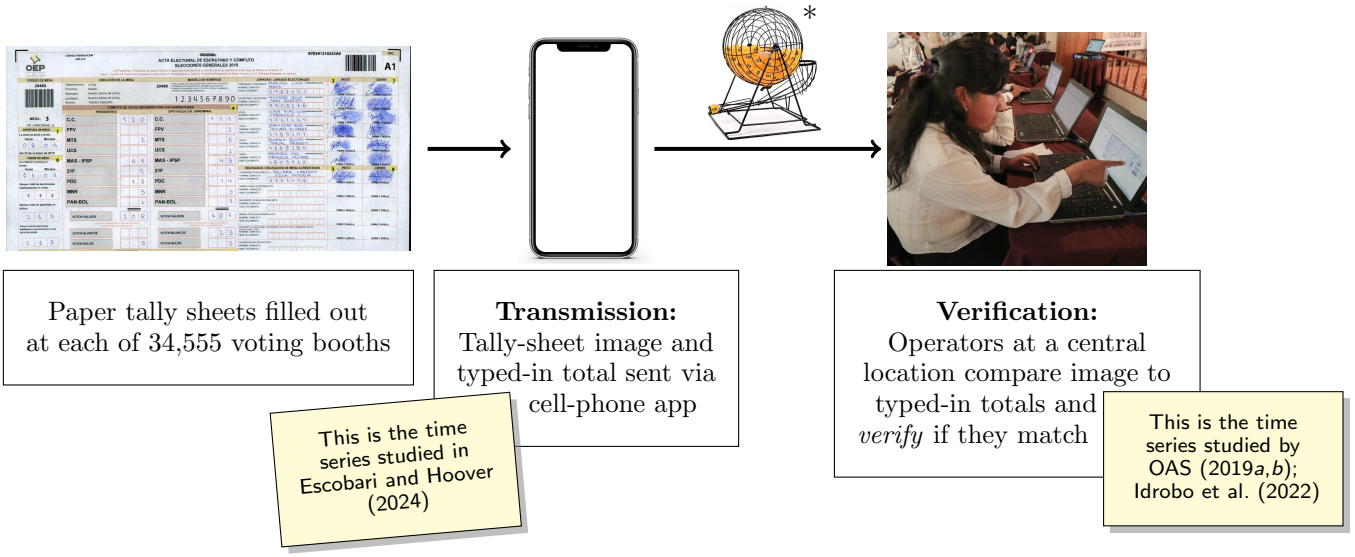
Background. In October of 2019, Evo Morales—a left populist then in his thirteenth year as President of Bolivia—competed for reelection against eight challengers. The fact of Morales's candidacy was itself controversial: the Bolivian constitution forbade another term in office; in a referendum, voters had previously defeated a proposal to lift term limits; and Morales was only able to run at all because a court ruled that term limits violated his human rights. Even friendly observers doubted Morales's commitment to certain democratic principles (Anria, 2016). And polls predicted that Morales's margin over the runner-up might or might not clear the bar needed to avoid a runoff: ten percentage points. Which is all to say that, even if everything had gone smoothly, election night was bound to be tense.

Everything did not go smoothly. The details of what went wrong are essential to understanding Escobari and Hoover's analysis, and instructive about the relationship between technical errors and perceptions of electoral integrity more generally (on this topic, see Antenangeli and Cantú, 2019).

Bolivia has an extremely fast system for aggregating voters' paper ballots into a

Figure 1: Bolivia's Preliminary Vote-Counting System

This figure summarizes the process by which each voting booth's tally is added to the preliminary vote count. The *transmission* and *verification* time series are uncorrelated during "the shutdown;" previous work studies the latter, while Escobari and Hoover focus on the former.



∗ From the pool of tally sheets transmitted through the cell-phone app before any given point in time, sheets are drawn for verification in an order that is effectively random (Distant Comment, 2023).

preliminary count.[1] Poll workers at each voting booth[2] count the paper ballots, fill out a physical paper tally sheet, type the vote totals into a mobile app, take a photo of the completed tally sheet, and then transmit both the typed-in totals and the photo through the app to the electoral authority. Operators at a central location then compare each tally sheet image to the vote totals typed in to the app; if they match, the tally sheet is *verified* and added to the preliminary count. Figure 1 illustrates this process. On the evening of the 2019 presidential election, this system counted more than 80% of tally sheets within hours of the polls closing.

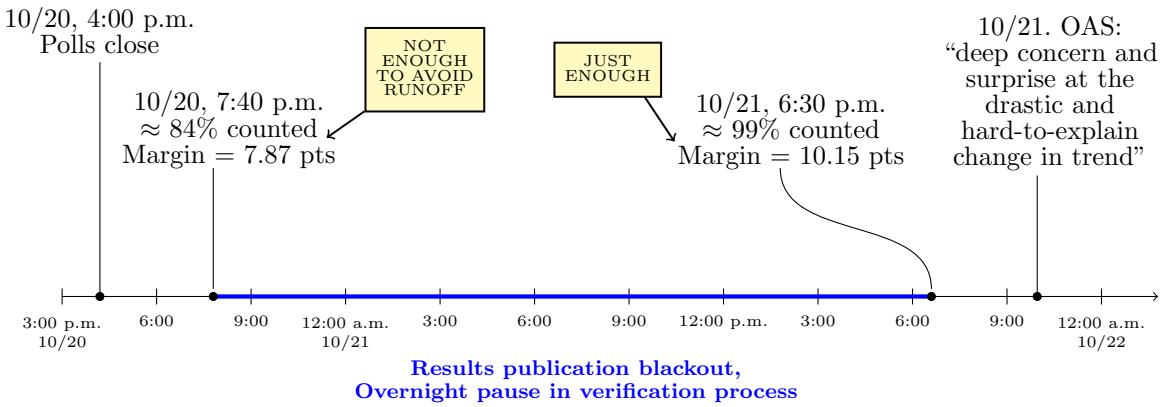
The process of tally-sheet verification was meant to proceed continuously overnight. Instead, verification abruptly paused at 8:07 p.m. on election night, and it did not

[1]There is also a definitive (rather than preliminary) count, in which the paper ballots are physically transported to regional offices of the electoral authority; this arrival of paper tally sheets is not the focus of Escobari and Hoover's analysis. See Idrobo, Kronick and Rodríguez (2022)for additional information about the two systems.

[2]There are several voting booths within each polling place. We use the term *voting booth* as a translation of *mesa*, or table, even though many *mesas* are not physical booths. Other work on this topic translates *mesa* as *polling station*, a term that we believe is too easily confused with *polling place*.

Figure 2: Understanding "The Shutdown"

This timeline of events surrounding the Bolivian presidential election on October 20, 2019, highlights the period referred to as "the shutdown."



resume until the following morning. Years later, Bolivian prosecutors—with the assistance of international experts—determined that the interruption stemmed from a series of technical errors (Atahuichi 2021, for a detailed and documented narrative, see Why the TREP Was Suspended 2023). But the arcane details of these technical errors were not immediately apparent. And because of the overnight pause in tally-sheet verification and the associated chaos within the electoral authority, the government effectively implemented a public information blackout: from 7:50 p.m. on election night until 6:30 p.m. the next day, the government provided no information about the progress of the preliminary count, leaving voters and electoral observers waiting in vain for updates. A website meant to provide real-time vote totals instead stayed frozen. Electoral officials held no press conferences. This silence aroused suspicion, as information blackouts often do (Antenangeli and Cantú, 2019).

When electoral authorities finally broke their silence, on the evening of the day after voting, they announced a razor-thin Morales victory: his margin had just barely exceeded the ten percentage points that he needed to avoid a runoff (see Figure 2 for a timeline). The Organization of American States (OAS) joined a chorus of voices crying fraud. The subsequent statements, audit, and reports of the OAS convinced many observers that the Morales administration had, in fact, stolen the election (Crisis Group, 2020). On the day that the OAS released its final audit report, the Bolivian military asked Morales to resign. He complied and fled to Mexico.

In the four years since that day, researchers have revisited the claims made in the

influential OAS reports.[3] Escobari and Hoover take a different tack. Rather than reconsider the analysis of the OAS, they instead highlight patterns in the data that were *not* identified in the OAS reports, claiming that these new patterns are also indicative of electoral fraud.

In particular, the OAS reports focus on how vote shares changed as a function of the time at which employees of the electoral authority added each voting booth's tallies to the preliminary count (OAS 2019*a,b*; Idrobo, Kronick and Rodríguez 2022), or *verification time*. This was the process that paused overnight. Escobari and Hoover instead turn attention to *transmission time*, i.e., the time at which poll workers at each electoral precinct transmitted results through the mobile app to the electoral authority (see Figure 1);[4] Escobari and Hoover refer to this time series as "arrival." The transmission process continued smoothly throughout election night; it did not pause during the shutdown. As we explain in more detail in Idrobo, Kronick and Rodríguez (2022)(Appendix A), transmission time and verification time are not at all the same. We should therefore think of Escobari and Hoover's analysis as an entirely new approach to evaluating allegations of fraud in this election, rather than as a replication or extension of the influential analysis of the Organization of American States (November 10, 2019*b*; December 4, 2019*a*) or of our previous work (Idrobo, Kronick and Rodríguez, 2022).[5]

Escobari and Hoover identify three allegedly suspicious patterns in the transmission time series. In what follows, we describe these patterns, note why Escobari and Hoover claim that they indicate fraud, and then explain why we do not view these patterns as indicative of fraud.

1. A difference-in-differences with converging pre-trends. In Escobari and Hoover's first empirical exercise, they consider a difference-in-differences. The difference-in-differences compares vote margins across two elections—2019, the contested elec-

[3] Johnston and Rosnick (2020); Rosnick (2020*a,b*); Nooruddin (2020*a,b,c*); Idrobo, Kronick and Rodríguez 2022.

[4] The physical paper tally sheets are also transported to regional offices of the electoral authority, as part of the definitive (rather than preliminary) count; this arrival of paper tally sheets is not the focus of Escobari and Hoover's analysis. See Idrobo, Kronick and Rodríguez (2022) for additional information about the two systems.

[5] Part of Escobari and Hoover's analysis comments directly on our analysis of within-precinct time trends in vote share. That part of Escobari and Hoover's paper appeared in their working paper, and we addressed their comments in Appendix D of Idrobo, Kronick and Rodríguez (2022). On this particular point, there is nothing new in the published version that merits additional discussion.

tion, and 2016, the previous poll—before and after the shutdown.[6] The implicit hypothesis underlying this analysis is that, in 2019, Bolivia's electoral authority tampered with the results in voting booths that were verified after the shutdown. The implicit assumption underlying the analysis is that, absent fraud, results in "treated" voting booths (those verified after the shutdown) would follow the time trend of the results from 2016, when Bolivians voted against Morales's proposal to lift term limits. The 2016 referendum was not marred by accusations of fraud; in fact, international observers described substantial improvements in electoral administration (in part due to the introduction of biometric voter ID) and "public opinion of electoral integrity was strengthened" (Cantú, 2023, 120).

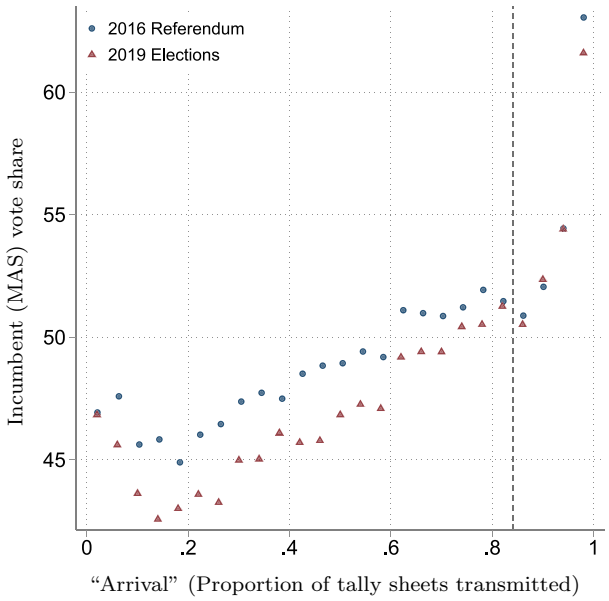
Escobari and Hoover present a graph that allows readers to see the time trends in vote share in both elections. We reprint this graph in Figure 3. For well-understood reasons, there is a "blue shift" in Bolivia just as there is a "blue shift" in the United States: Morales's vote share, like the vote share of the Democratic party in many U.S. states, tends to increase as the vote count progresses (incidentally, blue is also the color of Morales's political party, the MAS). This overall trend, which is evident in both elections, is a predictable consequence of electoral administration in Bolivia. Voters and poll workers with higher levels of education typically complete the voting and counting process marginally faster. Counting hundreds of paper ballots by hand, deciding which are valid and which are null, and filling out the many blank fields on tally sheets are all tasks that become faster with full literacy and numeracy. And because so many voting booths report in such a tiny window of time, small differences in reporting time can make a big difference in reporting percentile: at the busiest moment in the reporting window, a delay of just ten minutes moves a voting booth from the 36th to the 46th percentile of transmission time. Education is, of course, negatively correlated with support for Morales. That is why more pro-Morales voting booths tend to transmit results somewhat later in the count.

What appears concerning to Escobari and Hoover is not the mere fact of a pro-Morales time trend in vote share, but rather the difference-in-differences: Morales's 2019 vote share is closer to his 2016 vote share *during* the shutdown than *before* the shutdown, as is evident from Figure 3.

[6]Note: As explained in the previous section, Escobari and Hoover's difference-in-differences analysis considers trends in *transmission* time; the graph plots *transmission* time on the x-axis. But the treatment indicator is defined as a function of verification time.

Figure 3: A Difference-in-Differences with Converging Pretrends

Escobari and Hoover present this graph (and associated regression output) as evidence in support of two claims: (1) that the difference-in-differences is indicative of fraud in 2019 in the *post* period, and (2) that the converging pre-trends themselves constitute evidence that fraud was "a treatment that grows gradually" because "it takes time to implement fraud."



Escobari and Hoover's main regression specification estimates the difference-in-differences illustrated in Figure 3, controlling for "last names, geography, socioeconomic characteristics, and voting preferences."[7] Their estimate is positive, meaning that Morales's 2019 vote share is closer to his 2016 vote share *during* the shutdown than *before* the shutdown (consistent with the raw data in Figure 3). Escobari and Hoover interpret this difference-in-differences as "the exogenous effect of the fraud treatment" (p. 2). Absent fraud, they suggest, the incumbent's 2019 vote share would have been no closer to his 2016 vote share *during* the shutdown than *before* the shutdown. One problem with this interpretation, in our view, is that the pre-trends are clearly converging: 2019 vote share draws closer to 2016 vote share throughout most of the transmission window, not just during the shutdown (Figure 3). The pre-shutdown trends alone would therefore lead us to expect the 2019 vote share to rise closer to

[7]We cannot replicate their results because the authors declined to provide code or the final analysis data set, sharing only the source data files. Rebuilding their analysis data set from source data requires: (1) a merge on more than 5,000 precinct names, (2) deciding how to treat precincts that appeared or disappeared between 2016 and 2019, and (3) within matched precincts, re-weighting voting booths (*mesas*) when the number of voting booths changes, among many other researcher choices. Escobari and Hoover do not specify how they address any of these challenges.

2016 vote share during the shutdown.

To address the issue of converging pre-trends, Escobari and Hoover estimate a second specification in which they allow the two elections to follow different linear trends. This difference-in-differences is also positive: the incumbent earns a higher vote share during the shutdown than linear trends would lead us to expect (see their Table 3, p. 7); the authors interpret this result as evidence that "even after relaxing the common trends assumption, the shutdown has [a] statistically significant effect" (p. 6).

One problem with this interpretation, in our view, is that there is no reason to assume linearity in the secular election-specific trends. The authors show visually that the raw pre-trends are not parallel (Figure 3); they never evaluate—either via visual inspection or with formal tests—whether the pre-trends are in fact parallel after introducing global linear trends or other controls (on this topic, see Rambachan and Roth, 2023). In other words, Escobari and Hoover do not show that Morales's vote share in voting booths verified during the shutdown ("treated") was any higher than we would expect given pre-trends and observable characteristics.

2. The difference in trends itself. Escobari and Hoover then interpret the steeper 2019 time trend itself (relative to 2016) as evidence that fraud was "a treatment that grows gradually" because "it takes time to implement fraud." In other words, they pivot from evaluating the hypothesis of fraud *during the shutdown* to evaluating a different hypothesis of fraud throughout election night. This pivot entails implicitly asserting that, in the absence of fraud, the 2019 vote share would have followed the trend set by 2016 (again, see Figure 3). They estimate that, had the (linearized) 2019 time trend followed the (linearized) 2016 time trend, the incumbent's overall vote margin would have been 2.88 points lower. That this estimate emerges after including various controls "is in favor of interpreting the 2.88% as electoral fraud" (p. 7).

Our view is that there is no reason to expect that the trends in transmission time *should* be identical across elections. Transmission time is correlated with the geographic location of the polling place and with the socio-economic characteristics of the voters at each voting booth; any political shift that widens urban–rural cleavages, for example, or the political gap between more- and less-educated voters, would likely affect how the incumbent's vote share changes with transmission time. Many such shifts occurred between 2016 and 2019.

3. A discontinuous jump in vote share. Previous accusations of fraud in Bolivia's 2019 presidential election, including the influential accusations levied by the Organization of American States, posit a hypothesis of centralized tampering: that employees of the electoral authority in La Paz added votes for the incumbent after local poll workers had counted and reported results. Studying the *verification* time series (see Figure 1), the preliminary OAS audit report (November 10, 2019*b*) highlighted an apparently anomalous discontinuous jump in vote share when comparing tally sheets *verified* just before vs. just after the overnight puase in the verification process; in our previous work and in the following sections, we explain why these differences are not indicative of fraud (in short, tally sheets are not verified in the order in which they are transmitted; see Idrobo, Kronick and Rodríguez 2022, Appendix A).[8]

Escobari and Hoover, in contrast, propose a hypothesis of decentralized tampering. In their account, poll workers at thousands of electoral precincts across the country added votes for the incumbent and/or subtracted votes for the opposition, *before* reporting those totals to the central electoral authority. Precincts where such fraud took place would likely report later in the count for two reasons, according to Escobari and Hoover. First, only late in the evening did it become clear how close the election would be, perhaps creating new pressure for fraud. Second, "it takes time to implement fraud" (for example, "to rewrite minutes and forge signatures").

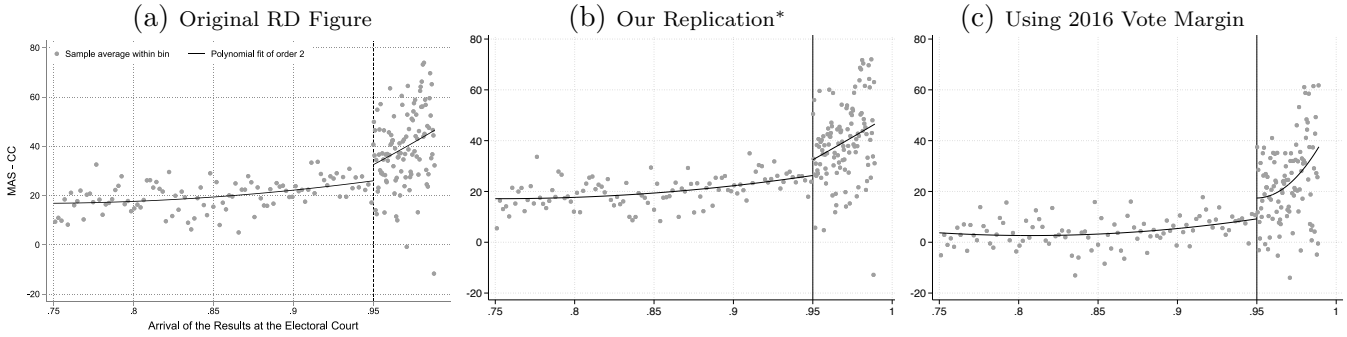
Escobari and Hoover explain how such decentralized tampering might create a discontinuous jump in vote share at specific *arrival* times (i.e. transmission times):

> "Discontinuities might simply be the result of fraud that takes place in different locations and that they coincide at the end, as pressure builds up to make sure the MAS-CC gap is big enough to avoid a runoff. ... [the] RD design isolates the treatment variation as a consequence of agents' inability to precisely control the assignment variable. In our case there is some control as fraudulent individuals can decide to delay the submission of forged minutes. However, they cannot go back in time. If a sufficiently large number of fraud polling stations build up close to the end, there is no point in delaying and a discontinuity might be unavoidable."

[8] Adding to the confusion, the *final* report of the OAS (December 4, 2019*a*) then studied a *different* cutoff in the verification time series where, as we explain in Idrobo, Kronick and Rodríguez 2022and below, there actually is no break.

Figure 4: An RD in which Placebo Outcomes Jump at the Boundary

Fig. (a), reprinted from Escobari and Hoover, reveals that there is a discontinuous jump in vote margin when 95% of tally sheets were transmitted (as we note in the main text, this is neither the running variable nor the cutoff studied in previous work). Escobari and Hoover present Fig. (a) as evidence of decentralized fraud, in which individual poll workers at thousands of voting booths across the country tampered with the vote and then coincidentally submitted their fraudulent tallies at the same moment. Rosnick (2022) observes—and we replicate in Fig. (c)—that 2016 vote shares also jump discontinuously at the same cutoff.



* As noted above (see Footnote 1 for details), we cannot exactly replicate the authors' results because they declined to provide replication code or their final analysis data set. That vote margin appears to have higher variance to the right of the cutoff is an artifact of the default approach to binning in `rdplot`, rather than a feature of these data.

In other words, even in the absence of coordination, local poll workers committing fraud might happen to transmit their falsified tally sheets through the mobile app all at the same time, creating a discontinuous jump in vote share.

This theory is imprecise. For one thing, Escobari and Hoover do not specify whether they suppose that (1) the government had previously instructed local poll workers to start falsifying tally sheets if the incumbent's margin was below a certain threshold at a certain time in the evening, or (2) local poll workers independently and of their own accord responded to news about the incumbent's margin by beginning to falsify tally sheets. Both possibilities strike us as implausible. In the former scenario, the subsequent OAS audit and/or local reporters would likely have uncovered evidence of such mass communication. The latter scenario, in contrast, cuts against the conventional wisdom about the relationship between vote margin and poll-worker motivation (which holds that losing demotivates poll workers). For another, Escobari and Hoover never explain why the coincidental simultaneity of the transmission of falsified tally sheets would occur when 95% of the voting booths had reported, rather than 91%, or 96%, or any other moment in the count. (We discuss the choice of 95% in the following section.)

Moreover, Escobari and Hoover's analysis violates one of the cardinal rules of regression discontinuity designs, which is that the cutoff is exogenously given rather than endogenously chosen by the researcher (Cattaneo, Idrobo and Titiunik, 2020, p. 5). Cattaneo and Titiunik (2024) are very clear on this point, writing that "The RD design can only be invoked if an RD treatment assignment rule actually occurred ... the RD design exists independently of the researcher, and can be verified externally." There is no sense in which Escobari and Hoover's RD fulfills this key condition. There is no sense in which "the assignment of treatment follows a rule that is known and hence empirically verifiable" (Cattaneo, Idrobo and Titiunik, 2020).[9]

Even if we overlook this major violation, however, and even if we grant that decentralized fraud could in theory produce a discontinuous jump in vote share at an arbitrary transmission time, the converse does not hold: the presence of a discontinuous jump in vote share at an arbitrary transmission time does not imply the existence of decentralized fraud. Nor is fraud the most likely explanation. Indeed, Rosnick and Wu (2022) observe—and we replicate in Figure 4—that 2016 vote shares also jump discontinuously at the same cutoff.[10] In other words, vote shares in an election that Escobari and Hoover themselves present as a clean control case follow the same pattern that they interpret as indicative of fraud in 2019. This result (the presence of a discontinuous jump in 2016 vote share at the same cutoff) reveals instead that underlying, pre-existing political attitudes are not smooth at the boundary.

Discussion. The findings that we discuss here are but three of many claims of statistical evidence of fraud in the 2019 Bolivian presidential election. To put the discussion in context, we summarize Escobari and Hoover's analysis alongside previous claims in Table 1. Following Eggers, Garro and Grimmer (2021, p. 6), we first ask whether the allegedly anomalous fact is a fact; if so, we ask whether it is anomalous. Synthesizing the allegedly anomalous statistical facts in this way spotlights how one type of error can lead to another.

The first high-profile quantitative analysis of Bolivia's 2019 election returns appeared

[9] In the influential regression discontinuity analysis in the preliminary report of the OAS (which, again, used *verification time* rather than *transmission time* as the running variable), the cutoff was indeed exogenously given: it was the time at which verification paused overnight (November 10, 2019*b*). In the final report of the OAS (December 4, 2019*a*), the cutoff was endogenously chosen, as in Escobari and Hoover.

[10] As noted in Footnote 1 above, we cannot exactly replicate Escobari and Hoover's results because they declined to provide code or their final analysis data set.

in the preliminary report of the Organization of American States (OAS) electoral audit, published on the morning of Sunday, November 10, twenty-one days after the election (OAS, November 10, 2019*b*). Many observers, including the OAS, had long since accused the government of foul play. But the November 10 report was the first to back these accusations with quantitative analysis: specifically, a figure. The figure plotted Morales's vote share against the proportion of the vote *verified* in the preliminary results system, marking with a vertical line the cutoff dividing tally sheets verified before the shutdown (that is, tally sheets verified before 8:07 p.m. on election night) from those verified after the shutdown (when verification resumed the following day). This figure revealed a discontinuous jump at the cutoff: Morales's vote share was *much* higher in the first tally sheets verified after the shutdown than in the last tally sheets verified before the shutdown. The clear implication was that, during the overnight shutdown, the government had tampered with the software in such a way as to add votes for Morales (and/or subtract votes for opponents) in order to ensure that he cleared the ten-percentage-point margin that he needed to avoid a runoff. At first glance, this graph appeared convincing: what else could explain such a large and discontinuous jump in the time-trend of Morales's vote share? Evo Morales himself seems to have found it alarming; he responded by immediately calling for new elections.

This allegedly anomalous statistical fact—that Morales's vote share jumps discontinuously between the pre- and post-shutdown tally sheets—is indeed a fact. The discontinuity at the shutdown is not the artifact of a methods or coding error. Yet it is not actually anomalous. Rather, this discontinuity is a predictable result of Bolivia's preliminary vote-counting system. After tally sheets are *transmitted* through the mobile app to the electoral authority, they are then drawn for verification in an order that is effectively random (Distant Comment, 2023). At times in the count when verification kept pace with transmission, verification order largely preserved transmission order (because there were few tally sheets waiting for verification at any given time). But during the shutdown, transmission continued even as verification paused for more than twelve hours. When verification resumed the next day, tally sheets were drawn in an effectively random order from the large pool of tally sheets that had been transmitted overnight. Given the overall upward trend in Morales's vote share, it is then unsurprising that his vote share was much higher in tally sheets verified just after the shutdown than in tally sheets verified just before the shutdown.[11]

[11]To see this, consider the following example (from Idrobo, Kronick and Rodríguez (2022)):

The *preliminary* report of the OAS audit thus reported an actual statistical fact and misinterpreted it as anomalous rather than innocuous. The *final* report of the OAS—for which the statistical analysis was conducted by an outside contractor rather than internal analysts—committed different errors. The final report presented two allegedly anomalous statistical facts that, it turned out, were not facts.

First, the final report claimed that Morales's vote share jumped discontinuously when 95% of the vote had been verified in the preliminary results system, a cutoff that *almost* (but not quite) corresponds to the shutdown cutoff of 95.6%. It is not clear why the preliminary report of the OAS examined the exogenous cutoff of 95.6% while the final report then switched to examining an arbitrary cutoff of 95%, nor is it clear that a discontinuous jump at 95% would be indicative of electoral fraud. But in any case, there is no discontinuous jump in vote share at 95% of the vote verified Idrobo, Kronick and Rodríguez (2022); the appearance of a jump at 95% was the artifact of a methodological error.

Second, the final report also examined an entirely different time series (i.e., a different running variable): the time at which each tally sheet was counted in Bolivia's slow, deliberate, paper-based definitive results system, which is separate from the preliminary results system. For an unstated reason, the final report of the OAS studied a cutoff of 95% of the vote counted in the definitive results system; this cutoff does not correspond to the shutdown or to any other known treatment, nor does it correspond to 95% of votes verified in the preliminary results system. In any case, the final report of the OAS presented a graph in which it appeared that Morales's vote share increased sharply (though continuously) in the final 5% of the definitive count. It is unclear whether this pattern would be anomalous if it existed. But in any case, there is no such sharp increase in Morales's vote share in the last 5% of the definitive count: that pattern was the artifact of a coding error. The time stamps were sorted alphabetically (with 7:01 p.m. following 7:01 a.m.) rather than chronologically.

"Imagine that there are 100 tally sheets and that the incumbent's vote share increases in transmission order, at a constant rate: he earns, say, 10% of votes on the tally sheet transmitted first, 10.5% of votes in the tally sheet transmitted second, 11% on the tally sheet transmitted third, and so on, until earning 60% of votes on the tally sheet transmitted last. Now imagine that tally sheets 1–50 are verified in the order in which they were received but that tally sheets 51–100 are verified in a random order (because verification paused between tally sheet 50 and tally sheet 51). If we plot the incumbent's vote share against the verification order, we should expect a jump between the fiftieth verified sheet (vote share = 25%) and the fifty-first verified sheet, which is drawn at random from the remaining tally sheets (expected vote share = 42.5%)."

The sequence of cutoffs studied in the influential OAS reports suggests a special problem with the RD part of Escobari and Hoover's analysis. The preliminary OAS report studied a cutoff of 95.6% of votes verified in the preliminary results system (which corresponds to the shutdown); the final report studied two additional cutoffs: 95% of votes verified and 95% of votes counted in the definitive, paper-based results system. Escobari and Hoover's choice of a fourth cutoff for their regression discontinuity analysis—95% of votes transmitted to the preliminary results system—bears a superficial resemblance to the cutoffs studied by the OAS. A casual reader might therefore skim Escobari and Hoover's analysis and glean that there *was* in fact a discontinuous jump in vote share at 95% after all, drawing the incorrect conclusion that the much-disputed OAS analysis was right all along.

Conclusion. In evaluating statistical claims of fraud in the 2020 presidential election in the United States, Eggers, Garro and Grimmer (2021) find that, for each claim, "what is purported to be an anomalous fact about the election result is either not a fact or not anomalous." Our reanalysis of Escobari and Hoover reaches the same destination. None of the facts that they establish—that the difference between late- and early-counted booths is greater in 2019 than in the previous election (2016), that the gap between 2019 and 2016 vote shares closes over the course of election night, or that the incumbent's vote margin jumps discontinuously at an arbitrary moment in the count—are actually anomalous. None of these facts indicates that the government of Evo Morales (then the incumbent) committed electoral fraud.

Settling this debate informs the public understanding of a pivotal moment in the history of the Americas. If there is no statistical evidence that the Morales government committed electoral fraud, then some might say that the Organization of American States leveled unfounded accusations that led to a coup d'etat in which an elected administration, however flawed, was replaced by a violent right-wing regime (Grandin, 2019). If, on the other hand, there *is* statistical evidence that the Morales government committed electoral fraud, then perhaps his ouster was less a coup than a stirring example of people defending their democracy (Mounk, 2019), with the OAS exerting characteristically pro-democratic "Western leverage" (Levitsky and Way, 2010). Which of these narratives is closer to correct bears on the story of the whole region in the twenty-first century. That is one reason that statistical minutia about this election are so hotly contested.

Similar conflicts inspire ad hoc election forensics all around the world. In the United States, in Honduras, in South Korea, in Mexico, and in many other places, researchers and reporters and hobbyists analyze electoral returns and interpret the results. Statisticians, economists, and political scientists have worked to develop methods for investigating fraud based on patterns in electoral returns (Alvarez, Hall and Hyde, 2009; Myagkov, Ordeshook and Shakin, 2009; Hicken and Mebane, 2017; Eggers, Garro and Grimmer, 2021; Grimmer and Ramaswamy, 2024). It is by following these established practices, rather than through improvised attempts to identify and ex-post interpret seeming statistical anomalies, that analysts can best uncover the truth about electoral integrity.

Table 1: Allegedly Anomalous Statistical Facts about Electoral Fraud in Bolivia in 2019

Organization of American States, Preliminary report (OAS, November 10, 2019*b*)		
Allegedly anomalous fact	Is it a fact?	Is it anomalous?
Morales's vote share is discontinuously higher in the tally sheets verified just after the shutdown than in tally sheets verified just before the shutdown. The cutoff, which is exogenously given by the shutdown time, occurs when 95.6% of votes were verified.	Yes	No. Because tally sheets are drawn for verification in an effectively random order from the pool of already-transmitted tally sheets, and because of the overall upward trend in Morales's vote share, this discontinuous jump is precisely what we would expect in the absence of fraud.

Organization of American States, Final audit report (OAS, December 4, 2019*a*)		
Allegedly anomalous fact	Is it a fact?	Anomalous?
Morales's vote share jumps discontinuously after 95% of votes were verified.	No. Unlike the jump at 95.6%, the apparent jump at 95% was an artifact of using an inappropriate estimator.	N/A
Morales's vote share jumps discontinuously after 95% of votes were counted in the definitive, paper-based system.	No. This apparent discontinuity was an artifact of a coding error: sorting the time stamps alphabetically (7:01 a.m., 7:01 p.m.) rather than chronologically.	N/A

Escobari and Hoover (2024)		
Allegedly anomalous fact	Is it a fact?	Is it anomalous?
Morales's vote 2019 vote share is closer to his 2016 vote share in voting booths verified after the shutdown than in voting booths verified before the shutdown.	Yes	No. As Figure 3 makes clear, the gap between Morales's 2019 and 2016 vote shares narrows throughout the count (i.e., even before the shutdown)—not just after the shutdown.
The time trend in Morales's vote share in 2019 is different from the time trend in vote share in 2016.	Yes	No. Any political shift that affects urban–rural cleavages, for example, or the gap between more- and less-educated voters, would change the time trend in vote share. Many such shifts occurred between 2016 and 2019.
Morales's vote share jumps discontinuously after 95% of tally sheets were transmitted.	Yes	No. The cutoff was endogenously chosen by the researchers and corresponds to no known event or treatment. Moreover, placebo outcomes are not smooth at the cutoff.
Morales's vote share increases with transmission time even within precincts, even after controlling for correlates of socio-economic status.	Yes	No. Differences in education across voting booths within the same precinct *should* create a within-precinct correlation between transmission time and vote share.

References

Alvarez, R.M., T.E. Hall and S.D. Hyde. 2009. *Election Fraud: Detecting and Deterring Electoral Manipulation*. Brookings Series on Election Administration and Reform Brookings Institution Press.

Anria, Santiago. 2016. "More inclusion, less liberalism in Bolivia." *J. Democracy* 27:99.

Antenangeli, Leonardo and Francisco Cantú. 2019. "Right on Time: An Electoral Audit for the Publication of Vote Results." *Statistics, Politics and Policy* 10(2):137–186.

Atahuichi, Rubén. 2021. "La Fiscalía cierra el caso Fraude con una pericia internacional que no halló manipulación en 2019." La Razón.

Cantú, Francisco. 2023. "Un análisis forense de los referendos presidenciales en Bolivia de 2008 y 2016." *Revista Elecciones* 22(25):111–150.

Cattaneo, Matias D., Nicolás Idrobo and Rocío Titiunik. 2020. *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge University Press.

Cattaneo, Matias D and Rocio Titiunik. 2024. "Protocols for Observational Studies: An Application to Regression Discontinuity Designs." *arXiv preprint arXiv:2402.11640* .

Crisis Group, International. 2020. "Bolivia Faces New Polls in Shadow of Fraud Row." Crisis Group Latin America Briefing N °43.

Distant Comment. 2023. "How Did Bolivia's 2019 TREP Count Tally Sheets?" Github.

Eggers, Andrew C., Haritz Garro and Justin Grimmer. 2021. "No evidence for systematic voter fraud: A guide to statistical claims about the 2020 election." *Proceedings of the National Academy of Sciences* 118(45).

Escobari, Diego and Gary A Hoover. 2024. "Late-arriving votes and electoral fraud: A natural experiment and regression discontinuity evidence from Bolivia." *World Development* 173:106407.

Grandin, Greg. 2019. "What 'The New York Times' Got Wrong on Bolivia." The Nation.

Grimmer, Justin and Abhinav Ramaswamy. 2024. "An Evaluation of Fraud Claims from the 2020 Trump Election Contests." Working Paper.

Hicken, Allen and Walter Mebane. 2017. "A Guide to Elections Forensics: Research and Innovation Grants Working Papers Series." USAID.

Why the TREP Was Suspended. 2023. "Why The TREP Was Suspended." Reddit.

Idrobo, Nicolás, Dorothy Kronick and Francisco Rodríguez. 2022. "Do shifts in late-counted votes signal fraud? Evidence from Bolivia." *The Journal of Politics* 84(4):2202–2215.

Johnston, Jake and David Rosnick. 2020. "Observing the Observers: The OAS in the 2019 Bolivian Elections." Center for Economic and Policy Research.

Levitsky, Steven and Lucan A Way. 2010. *Competitive authoritarianism: Hybrid regimes after the Cold War*. Cambridge University Press.

Mounk, Yascha. 2019. "Evo Morales Finally Went Too Far for Bolivia." The Atlantic.

Myagkov, M., P.C. Ordeshook and D. Shakin. 2009. *The Forensics of Election Fraud: Russia and Ukraine*. Cambridge University Press.

Nooruddin, Irfan. 2020*a*. "Comment on the 2019 Bolivia Presidential Election and OAS Statistical Analysis.".

Nooruddin, Irfan. 2020*b*. "Replication Data for: OAS Audit of Bolivia General Election 2019." Harvard Dataverse.

Nooruddin, Irfan. 2020*c*. "Yes, Bolivia's 2019 election was problematic. Here's why." Washington Post Monkey Cage.

OAS. December 4, 2019*a*. "Electoral Integrity Analysis: General Elections in the Plurinational State of Bolivia, October 20, 2019, FINAL REPORT." Organization for American States (OAS), Secretariat for Strenthening Democracy (SSD), Department of Electoral Cooperation and Observation (DECO).

OAS. November 10, 2019*b*. "Electoral integrity analysis: General Elections in the Plurinational State of Bolivia, PRELIMINARY FINDINGS." Organization of American States.

Rambachan, Ashesh and Jonathan Roth. 2023. "A more credible approach to parallel trends." *Review of Economic Studies* p. rdad018.

Rosnick, David. 2020*a*. "The Ends Don't Justify the Means: Yet Another Analysis Fails to Support the OAS' Delegitimization of Bolivian Elections." Center for Economic and Policy Research.

Rosnick, David. 2020*b*. "The OAS Accusation of Electoral Fraud Against Evo Morales Is Bullshit — And Now We Have the Data to Prove It." *Jacobin* .

Rosnick, David and Aileen Wu. 2022. "Nickels Before Dimes." CEPR.