Do Shifts in Late-Counted Votes Signal Fraud? Evidence From Bolivia

Idrobo, Nicolás and Kronick, Dorothy and Rodriguez, Francisco

University of Pennsylvania, University of Pennsylvania, University of Notre Dame

September 2020

Do shifts in late-counted votes signal fraud? Evidence from Bolivia[†]

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

September 30, 2020

Abstract

Surprising trends in late-counted votes can spark conflict. When late-counted votes led to a narrow incumbent victory in Bolivia last year, fraud accusations followed—with dramatic political consequences. We study the pro-incumbent shift in vote share as the tally progressed, finding that we can explain it without invoking fraud. Two observable characteristics, rurality and region, account for most of the trend. And what looked like a late-breaking surge in the incumbent's vote share—which electoral observers presented as evidence of foul play—was actually an artifact of methodological and coding errors. Our findings underscore the importance of documenting innocuous explanations for differences between early- and late-counted votes.

[†] We are grateful to Santiago Anria and Marc Meredith for detailed guidance. For research assistance, we thank Mateo Arbeláez, Estefanía Bolívar Méndez, and Juan Vera. For comments, we thank María Eugenia Boza, Matias D. Cattaneo, Tulia Falleti, Camilo García-Jimeno, Etan Green, Guy Grossman, David Hausman, Dan Hopkins, Richard Kronick, Yphtach Lelkes, Matt Levendusky, Michele Margolis, Edward Miguel, Jonathan Mummolo, David Rosnick, Josh Simon, Uri Simonsohn, Tara Slough, and Rocío Titiunik. All errors are our own.

[‡] Department of Political Science, University of Pennsylvania.

[§] Corresponding author (kronick@upenn.edu). Dept. of Political Science, University of Pennsylvania.

[¶] Visiting Fellow, Kellogg Institute for International Studies, University of Notre Dame.

The order in which votes are counted is anything but random. Voters who have sat through election night in Brazil or Colombia or the United States, for example, know that results from the first few precincts rarely resemble the final outcome.[1]

Yet politicians often decry this fact as evidence of fraud. These accusations can spark conflict. In the 2007 presidential election in Kenya, for example, the opposition candidate led on election night but ultimately suffered a narrow loss (Kanyinga, 2009). His party protested. Hundreds were killed in the ensuing crisis; hundreds of thousands were displaced. The Kenyan case is extreme, of course. But even milder conflict over late-counted votes can induce harm.[2] Doubts about the legitimacy of the electoral process can demoralize and demobilize voters (Alvarez, Hall and Llewellyn, 2008; Birch, 2010; Simpser, 2012; Norris, 2014), even in the absence of actual electoral malpractice (Norris, Garnett and Grömping, 2020).

Documenting innocuous sources of shifts in late-counted votes may constrain politicians who would otherwise cry fraud. Scholars have done this work for the United States (Foley, 2013; Foley and Stewart, 2020; Li, Hyun and Alvarez, 2020; Cottrell, Herron and Westwood, 2018). But other countries' shifts in late-counted votes, while common, are poorly understood—leaving them open to politicized interpretation.

We revisit the controversial Bolivian presidential election of October 20, 2019. On election night, electoral authorities announced that incumbent Evo Morales held a 7.9-point lead over the runner-up—less than the 10 points needed to avoid a runoff. But the following evening, with nearly all of the vote counted, Morales's margin narrowly exceeded 10 points. The runner-up alleged fraud (Mesa, 2019). And critically, the Organization of American States (OAS) issued a statement expressing "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" (October 21, 2019*d*).

---

[1] One example from each case illustrates the point. Brazil: In the 2018 presidential election, Fernando Haddad earned just 27% of the first 93% of votes counted but more than 43% of the last 7%. Colombia: In the 2018 presidential election, Iván Duque earned a 13-point lead in the first 93% of the vote but just a 5-point lead in the last 7%. United States: In the 2018 congressional election, Young Kim (candidate for California's 39th district) held a 3-point lead in the with 65% of the vote counted—but ultimately lost by 3 points (Li, Hyun and Alvarez, 2020).

[2] There are many examples. In the 2017 gubernatorial election in Estado de México, the opposition candidate led in early-counted votes and declared victory—only to ultimately lose, by a narrow margin, to the candidate of the ruling PRI (Animal Político, August 8, 2017*a*). Though the reversal was likely due to late-counted votes from rural areas (Animal Político, June 5, 2017*c*), the opposition party challenged the result in court, and President Andrés Manuel López Obrador publicly questioned the integrity of the election (Animal Político, June 4, 2017*b*).

The political consequences were dramatic. In large part because of the fraud allegations, the Bolivian military asked Morales to resign; he fled to Mexico. An opposition-party senator took office as interim president. At this writing, she remains in office.

We study the pro-Morales shift in vote share as the tally progressed, finding that we can explain it without invoking fraud. Two observable characteristics—rurality and region—account for most of the trend; rural and highland precincts reported later, and rural and highland areas are known to favor Morales (Anria, 2018; Madrid, 2012).

Moreover, the pattern that the OAS presented as indicative of fraud—"a massive and unexplainable surge in the final 5% of the vote count" (OAS, 2019*a*, p. 94)—was actually an artifact of methodological and coding errors. First, when analyzing data from Bolivia's preliminary results system, the OAS used an inappropriate estimator to claim that there was a discontinuous jump in Morales's vote share when 95% of the vote had been counted (OAS, 2019*a*, p. 88). The jump does not exist; the incumbent's vote share is continuous at 95% of the count. Second, when analyzing data from Bolivia's definitive results system, the OAS sorted time stamps alphabetically, such that 7:01 p.m. comes right after 7:01 a.m.; when the time stamps are sorted chronologically, the apparent late-breaking increase in Morales's vote share disappears. This latter error was first noted in *Jacobin* magazine (Rosnick, 2020*b*), after the OAS consultant published replication data in response to an earlier draft of the present paper.[3] Throughout, we discuss other issues with the OAS's analysis and with follow-up studies (Escobari and Hoover, 2019; Newman, 2020).

The OAS's quantitative findings played an important role in the evolution of Bolivia's political crisis (Crisis Group, 2020, p. 3–4). The OAS drew an explicit connection between the quantitative findings and the outcome, stating that Morales's first-round victory was "only made possible by a massive and unexplainable surge in the final 5% of the vote count. Without that surge . . . he would not have crossed the 10% margin that is the threshold" (OAS, 2019*a*, p. 94). We find that there was no such surge.

Our analysis does not establish the absence of fraud; indeed, the OAS emphasized

---

[3] A note on sequencing: The OAS audit report was published in December, 2019. Johnston and Rosnick (2020) critiqued the report in great detail in March, 2020, but did not have access to the data used by the OAS. We obtained these data via the New York *Times* and posted the first draft of the present paper in June, 2020. At that point, the OAS had denied our request for replication materials. In response to our working paper, the OAS consultant published the replication materials in August, 2020 (Nooruddin, 2020*a*,*b*). We have since revised our paper to reflect what we learned from the replication materials.

other indicia of electoral malpractice, such as secret servers, falsified tally sheets, undisclosed late-night software modifications, and a fragile chain of custody for voter rolls and ballots, among other problems (OAS, 2019*a*).[4] We assess only the quantitative evidence, not the integrity of the election overall. Rather, we find that we do not *require* fraud in order to explain the quantitative patterns used to help indict Evo Morales. This finding underscores the importance of documenting innocuous explanations for differences between early- and late-counted votes.

We contribute to an ongoing debate over quantitative patterns in the Bolivian electoral returns.[5] Beyond Bolivia, we contribute to three literatures. First, our results echo work in American Politics about the "blue shift:" votes counted after election day disproportionately favor the Democrats (Foley, 2013; Foley and Stewart, 2020; Li, Hyun and Alvarez, 2020). While politicians and pundits often point to the blue shift as evidence of fraud, scholars find that it is predictable: young and non-white voters, disproportionately Democrats, are more likely to cast provisional and mail-in ballots, which are more likely to be counted late. In Bolivia, too, compositional changes likely explain the shift in late-counted votes.

Second, we contribute to literature on the role of international electoral observers (e.g. Donno, 2010, 2013; Hyde, 2007, 2011; Beaulieu and Hyde, 2009; Hyde and Marinov, 2014; Simpser and Donno, 2012; Bush and Prather, 2018; Kavakli and Kuhn, 2020). One central finding of previous work is that intergovernmental organizations (such as the OAS) are *less* likely to question electoral integrity than nongovernmental organizations (Kelley, 2009, 2012), perhaps because the former are beholden to member states, who may push for leniency. Indeed, in Kelley's data, the OAS itself—one of "a small core of organizations with a serious commitment to high-quality election observation" (Carothers, 1997, p. 21)—ranks among the observers least likely to criticize or condemn electoral integrity (2009, p. 779). In that sense, the Bolivian case constitutes something of an exception. On the other hand, the Bolivian case is consistent with Bush and Prather (2017), who find that third-party monitors can powerfully shape local perceptions of electoral credibility—especially those of political losers inclined to discredit the election anyway.

---

[4] Other authors claim that these findings do not reveal intentional electoral manipulation (Johnston and Rosnick, 2020). We restrict our analysis to the statistical evidence.

[5] OAS (2019*a*); Escobari and Hoover (2019); Johnston and Rosnick (2020); Williams and Curiel (2020); Mebane (2019); Nooruddin (2020*c*); Minoldo and Quiroga (2020); Newman (2020); Rosnick (2020*a*); Nooruddin (2020*a*,*b*); Rosnick (2020*b*).

Finally, our results highlight an opportunity for future work on electoral fraud. Political scientists have developed influential and sophisticated forensic tools for fraud detection (e.g. Hicken and Mebane, 2017; Alvarez, Hall and Hyde, 2009). We have fewer tools for assessing politicians' (often unsophisticated) claims of fraud. Like Goel et al. (2020), who debunk myths of widespread double-voting in the United States, our analysis takes a step in this direction. Future work could similarly apply econometric tools toward assessing controversial and consequential claims of fraud.

1 Context: Chronicle of a Crisis Foretold

On October 20, 2019, Bolivian voters cast ballots in the first round of a presidential election. The contest pitted incumbent Evo Morales against eight challengers. Morales, first elected in 2005 as part of Latin America's pink tide (Falleti and Parrado, 2018), was seeking a fourth term in office.

This alone was controversial. Bolivia's 2009 constitution imposed a two-term limit, but in 2013 courts had allowed Morales to run for a third term, on the grounds that his first term did not count because it began prior to the new constitution. In 2016, Morales held a referendum on his proposal to eliminate term limits all together—and voters defeated it, 51% to 49%.[6] Morales was able to run in 2019 only because Bolivia's highest court later ruled that term limits violated the American Convention on Human Rights (Anria and Cyr, 2019). The president of the electoral tribunal resigned in protest (Aguilar, 2018).

To avoid a runoff, Morales needed more than 40 percent of the vote and a 10-point margin over the second-place candidate (Bolivian Constitution, Article 166).[7] After the polls closed at 7:00 p.m., Bolivia's electoral authority began posting online results from the preliminary results system (see the following section for details on this system). At 7:40 p.m., the electoral authority initiated a planned pause in the public transmission of results, in advance of a scheduled press conference. The idea was to freeze website updates during the televised announcement, to avoid confusion (NEOTEC, October 28, 2019, p. 3). Just minutes earlier, the Panamanian cybersecurity company that the Bolivian government had hired to monitor the election issued a

---

[6] The two-term limit in the 2009 constitution was itself more favorable to the incumbent than the previous rule, which forbade immediate reelection, allowing reelection only after sitting out at least one term (Corrales, 2016, p. 8).

[7] Or an outright majority.

"maximum alert" about a burst of activity from one of the servers (Ethical Hacking, 2019, p. 35). At the press conference, which began at 7:50 p.m., authorities reported that, with 83% of voting booths reporting,[8] Morales had 45.71% of the vote to Carlos Mesa's 37.84%, a gap of 7.87 points (Bolivia tv).

Trouble began when the electoral authority did not resume the public transmission of the results. The reason is disputed. Critics charge that the government used the shutdown in order to tamper with the electoral results. The government claimed that they never intended to tally 100% of the vote in the preliminary results system (Los Tiempos, 2019*b*). Other accounts attribute the shutdown to an "enormity of technical fuck-ups" and "lack of expertise" (*impericia*) (Cambara Ferrufino, 2019).

Electoral authorities did not update the public results until the evening of the following day. By then, Morales had gained a 10.15% lead over Carlos Mesa (Los Tiempos, 2019*a*). Three days later, on October 24, the Plurinational Electoral Organ (OEP) published near-final results in which Morales won 47.05% to Mesas's 36.53%—a margin of 10.52 points, large enough for Morales to avoid a runoff.[9]

Opposition leaders cried fraud (AFP, 2019). Bolivia "exploded in protest" (Kurmaneav and Castillo, 2019); two protesters were killed, many were injured, offices of the electoral authority were vandalized, and a local MAS building was burned (MAS, for Movimiento al Socialismo, is Morales's party). Polls suggested that a run-off election would have been close, because opposition votes may have coalesced around Carlos Mesa (ANF, 2019).

Statements from the OAS played an important role in the evolution of Bolivia's political crisis (Crisis Group, 2020, p. 3–4). Together with the European Union, the OAS's Department of Electoral Cooperation and Observation sent a mission to observe the elections. On the evening of October 21, one day after the election, the mission issued a statement expressing "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," and "urg[ing] the electoral authority to firmly defend the will of the Bolivian citizenry" (OAS, October 21, 2019*d*). Two days later, on October

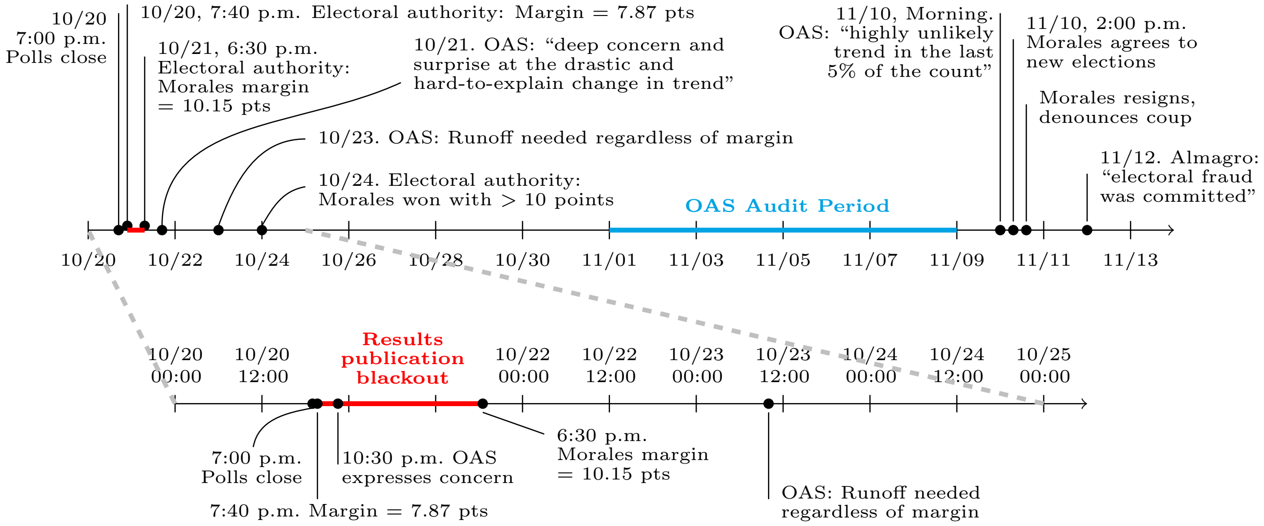
---

[8] The OAS later noted that 89%—not 83%—of tally sheets had been transmitted at this point, and that the electoral authorities "deliberately hid from citizens 6% of the tally sheets that were already in the [preliminary results system] but not published" (p. 4).

[9] The results announced on October 24 included more than 99.5% of the vote and thus Morales's first-round victory was irreversible. The final results, announced the following day (October 25), gave Morales 47.08% to Mesa's 36.51%, a margin of 10.57 points.

Figure 1: Key Announcements from the Bolivian Government and the OAS

For space reasons, we restrict the events on this timeline to the announcements key to understanding the statistical analysis. However, we note that protests began on election night and escalated as soon as electoral authorities announced the reversal the next evening.



23, the mission published a preliminary bulletin recommending that Bolivia hold a runoff election even if Morales were to earn a margin greater than ten points in the final tally (OAS, October 23, 2019*c*, p. 5). Figure 1 plots key announcements on a timeline. The OAS also called for all actors to abstain from violence.

Key political actors within Bolivia cited the OAS in calls for new elections and for Morales's resignation. For example, in a statement requesting the resignation of all electoral authorities and the convening of a new electoral process, Carlos Mesa's party summarized the OAS reports as "evidencing the violation of basic principles essential for the transparency of this electoral process and a sudden and inexplicable change of the irreversible trend towards a second round" (Comunidad Ciudadana, November 8, 2019). The opposition Committee for Santa Cruz even drafted a resignation letter for Morales and asked him to sign it; first on the Committee's list of reasons was the fact that "as the OAS delegate said, [the preliminary results transmission system] resumed with an inexplicable change in the vote trend" (CSC, November 4, 2019).

Amid continuing unrest over the disputed result, Morales's government signed an agreement with the OAS to conduct a formal audit (Flores and Valdez, 2019). The audit team published its preliminary report on the morning of Sunday, November

10. The report claimed to find evidence of secret servers, falsified tally sheets, and a deficient chain of custody for critical electoral material—as well as a “highly unlikely trend in the last 5% of the vote count” (OAS, November 10, 2019*b*, p. 9).

That afternoon, Morales responded by announcing that the government would convene new elections, under new electoral authorities (Collyns, 2019). But just hours later, under intense public pressure, Bolivia’s military chief and police chief asked Morales to resign (Kurmanaev, Machicao and Londoño, 2019). He stepped down that evening and flew to political asylum in Mexico, claiming that he had been ousted in a coup. Two days later, in a speech to the Permanent Council of the OAS, OAS Secretary General Luis Almagro said that “yes, there was a coup d’etat in Bolivia: it happened on October 20, when electoral fraud was committed.” In December, the OAS audit team published its final report (OAS, 2019*a*).

2 Data

Bolivian voters cast paper ballots at one of 34,555 voting booths (*mesas*) located within 5,296 precincts, or polling places (*recintos*). The ballot uses colors and photos as well as text to communicate voters’ choices (see Appendix Figure H.1 for an image).

Three types of poll workers administer the election. First, each voting booth has six “jurors” (*jurados*), who are (a) randomly selected from among each booth’s registered voters and (b) legally required to serve (Exeni Rodríguez, 2020). The jurors are responsible for checking voters’ names against the registration list, distributing and receiving ballots, and, most importantly, counting the paper ballots and writing the totals on a paper tally sheet (*acta*). Any citizen or party representative may observe this process. Second, an *electoral notary*, hired by the electoral authority, checks the tally sheet for obvious errors (TSE, 2019); there is one notary per precinct. Finally, a *preliminary results system operator*, also hired by the electoral authority, takes a photo of the tally sheet, transmits the photo to the electoral authority via an app, and types the vote totals into the app.

Two systems aggregate the tally sheets. The Transmission of Preliminary Electoral Results, or TREP (*Transmisión de Resultados Electorales Preliminares*), provides a preliminary count. After the preliminary results system operator transmits a tally sheet image and vote counts through the app, a team of verifiers look at the image and re-type the totals into the system. If these re-typed figures match the figures typed by the on-site operator, the tally sheet is recorded as *verified* and the numbers are added to the preliminary results system.

The second aggregation is the *cómputo*, or calculation, which is the legally binding official count. This count is much slower and more accurate than the preliminary results system. The paper tally sheets are delivered to electoral authorities in each of Bolivia's nine states (*departamentos*), where they are scanned and transmitted to the national electoral authority. Two separate teams independently transcribe the tally sheets. If the transcriptions match, the totals are added to the count (OEP, 2019, p. 5); otherwise, a third operator checks the transcription. These figures—not the preliminary results (TREP) numbers—determine the outcome of the election.[10]

Even though the preliminary results system time stamps include minutes and seconds, only 8% of tally sheets have unique time stamps. This makes sense given that there are 34,555 tally sheets, almost all of which arrived within two hours—7,200 seconds—of the polls closing.[11] For our analysis, we sort tally sheets in a random order within each time stamp.

3 Rurality and region explain most of the vote-share trend

Morales earned a much higher vote margin in late-reporting voting booths than in booths reporting early. The blue line in Figure 2 plots Morales's average margin over Civic Community as a function of reporting time; after declining at the very beginning of the count, Morales's average margin rose steadily through most of the active reporting window, from near zero to approximately 40 percentage points by the end of the evening.[12]

In the immediate aftermath of the election, several observers hypothesized that rurality might explain the strong pro-Morales trend in vote share. Greg Grandin, for

---

[10] In principle, the official *cómputo* count and the preliminary results system are separate; in practice, the OAS found evidence of contamination, with some preliminary figures funneled directly into the official count (OAS, 2019*a*, p. 6).

[11] Actually, as discussed in detail in the following section, only 33,038 tally sheets made it into the preliminary results system; the remaining 1,513 have no preliminary results system time stamps.

[12] The blue line in Figure 2 is fit only to observations in the estimation sample for Equation 1, and thus excludes precincts outside Bolivia (such as embassies). The time trend including all observations is quite similar; see Appendix Figure H.6.

example, suggested that "areas that reported later—in particular those that are more rural and/or poorer—are on average much more pro-Morales than the general electorate" (2019). This is intuitive, because rural-urban and regional cleavages structure Bolivian politics (Anria, 2018, Ch. 2). The workers, peasants, and indigenous groups of the eastern highlands constitute one regional bloc; the other is dominated by natural gas and big agriculture interests in the western lowlands. The highlands are much more supportive of Evo Morales and the MAS (Anria, 2018, p. 64–67). And within each regional bloc, MAS vote share increases both with rurality and with indigeneity (Madrid, 2012, p. 69). Poverty, illiteracy, and infant mortality remain higher in rural than urban areas, despite more than fifty years of convergence (Klein, 2011, p. 282).

To understand the extent to which changes in the regional and urban-rural composition of reported votes account for the trend in the blue line Figure 2, we estimate:

$$\text{MAS Margin}_{bpm} = f\left(\text{Time}_{bpm}\right) + \beta_2 \text{Lowlands}_{bpm} + \boldsymbol{\beta_3}\mathbf{Rural}_{bpm} + \epsilon_{bpm} \tag{1}$$

where MAS Margin_{bpm} is Morales's margin (-1 to 1) in voting booth b in precinct p in municipality m; Time_{bpm} is the time when voting booth b transmitted preliminary results; Lowlands_{bpm} is an indicator taking a value of 1 for voting booths in the lowland departments of Pando, Beni, Santa Cruz, and Tarija (also called the *medialuna* or half moon because they form a crescent around the highlands); and $\mathbf{Rural}_{bpm}$ is a vector of four correlates of rurality: (a) the (log) number of registered voters per square kilometer around precinct p, (b) the (log) proportion of the population employed in agriculture in municipality m, (b) the (log) population of municipality m, and (c) an indicator taking a value of 1 if municipality m is the capital of its department.[13] In some specifications (as noted), we also control for the average 2016 vote share in precinct p, i.e., $\overline{p}_{2016} = \frac{1}{|p|}\sum_{b\in p}(\text{MAS Margin 2016})_{bpm}$ (we can match precincts across elections, but not voting booths).

We estimate $f\,(\text{Time})$ using the semi-parametric estimator proposed by Robinson (1988).[14] Figure 2a plots the result. The blue line marks a non-parametric fit to

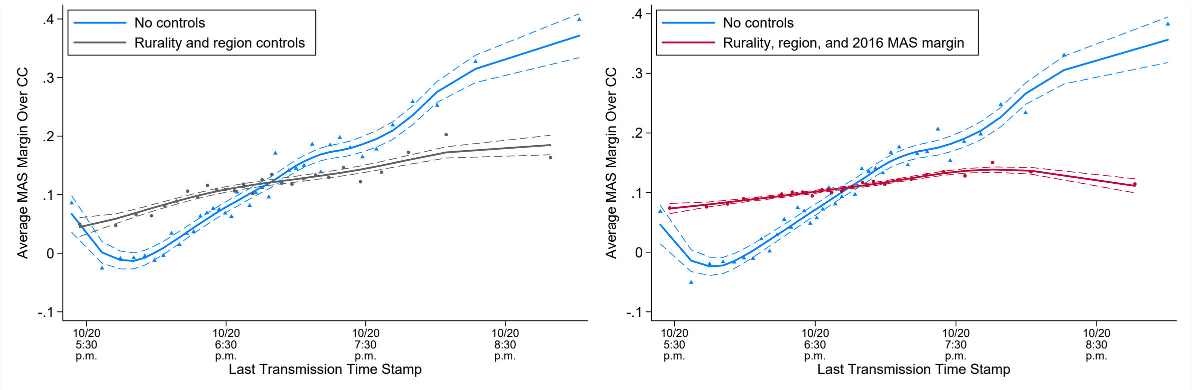
---

[13] The municipality-level measures come from the Bolivian census. The precinct-level density data were provided by a Bolivian PhD student who wished to remain anonymous. This researcher geocoded precincts, based largely on a publicly available geocoded data set of schools, and calculated registered voters / km^2 in 10km×10km grid cells.

[14] This analysis raises the question of what to do with a problematic set of observations: voting booths that never transmitted results through the preliminary system (for diverse reasons including lack of cellular service and tally sheet illegibility), and which therefore have no time stamps in the

Figure 2: Region and Rurality Explain Most of the Time Trend in Morales's Margin

The blue lines mark estimates of Morales's average margin over time. The gray line in (a) marks our estimate of f(Time) from Eq. 1, revealing that rurality and region account for much of the trend. The cranberry line in (b) marks our estimate of f(Time) from a specification that also controls for previous vote margin.



(a) Rurality and Region Controls

(b) + 2016 MAS Vote Margin Controls

Points mark the average MAS margin over Civic Community in optimal (data-driven) bins of the time variable (Cattaneo et al., 2019). The solid blue lines mark local linear fits following Calonico, Cattaneo and Farrell (2018); the samples for (a) and (b) are slightly different because we do not match all precincts to the 2016 data. The solid grey and cranberry lines mark estimates of f(Time) from Equation 1, using the semi-parametric estimator proposed in Robinson (1988). Dashed lines mark 95% confidence intervals. Both figures trim the top and bottom 2% of observations; for a version without trimming, see Appendix Figure C.1a.

the raw data. The grey line marks $\widehat{f}$(Time), our estimate of f from Equation 1. The comparison reveals that merely accounting for Bolivia's regional and urban-rural divides—even using crude measures—dramatically reduces the slope of the vote-share trend. (For visual clarity, we trim the top and bottom 2% of observations; Appendix Figure C.1a presents results without trimming. The takeaway is the same.)

Figure 2a also reveals that, controlling for region and rurality, the estimated relationship between Morales's margin and time ($\widehat{f}$) is approximately linear. This allows us to estimate a fully parametric version of Equation 1 in which $f\,(\text{Time}) = \beta_1\text{Time}$. We standardize the time variable so that a one-unit increase corresponds to a one-standard-deviation increase (approximately 45 minutes). Table 1 reports the results. When we omit the controls for region and rurality, $\widehat{\beta}_1 \approx 0.09$, which is to say that Morales's average margin (not cumulative margin) increases by 9 percentage points

---

preliminary results system (4.4% of observations). In the main text, we treat these observations as the latest reporters within their respective municipalities, assigning them the maximum of observed municipality reporting times. As we show in Appendix C, our results are not sensitive to other possible choices, such as assigning the minimum municipality reporting time, median municipality reporting time, or dropping these observations all together.

Table 1: Region and Rurality Alone Explain 2/3 of the Time Trend

Estimates of a fully parametric version of Equation 1, in which $f(\text{Time}) = \beta_1\text{Time}$. We standardize time so that a one-unit increase corresponds to one standard deviation, or $\approx$ 45 minutes. Column (1) presents the bivariate specification; Columns (2)–(5) sequentially add controls.

	(1) No controls	(2) $+\mathbb{1}$(Lowlands)	(3) +Rural	(4) +2016	(5) +Department +2016×Dep't
Reporting time (standardized)	0.089 (0.032)	0.063 (0.022)	0.031 (0.012)	0.015 (0.003)	0.005 (0.002)
Observations	31,529	31,529	31,529	29,241	29,241
R-squared	0.040	0.122	0.499	0.884	0.892

Standard errors, clustered at the municipality level, in parentheses. Column (2) includes an indicator for the lowland departments of Pando, Beni, Santa Cruz, and Tarija; Column (3) adds a vector of proxies for rurality; Column (4) adds average precinct MAS margin in the previous election (2016); Column (5) adds indicators for Bolivia's nine departments as well as interactions between each department indicator and the 2016 margin.

every 45 minutes (Column 1). Bolivia's regional and urban-rural divides account for two-thirds of this trend: when we include the region dummy (*Lowlands*) and proxies for rurality, the estimate of β_1 falls to 0.03. This is remarkable given the relative parsimony of Equation 1.

We also consider the previous poll, in 2016, when voters defeated Morales's proposed constitutional amendments in a yes-or-no referendum. That result was not contested; indeed, the OAS electoral observation mission made no reference to malpractice in its reports (2016*a*; 2016*b*).[15] When we control for the mean margin in each precinct in 2016 (we cannot match voting booths across elections), our estimate of β_1 falls another 50%, to 0.015 (Table 1, Column 4).[16] This is not an artifact of the linear approximation; in Figure 2b, we plot f (Time) from a semi-parametric specification in which we also control for the mean margin in each precinct in 2016.

When we include indicators for Bolivia's nine departments and interactions between those indicators and the 2016 precinct-level vote share (Column 5), β_1 falls an additional 66%, to 0.005, which is to say that every forty-five minutes Morales's average margin increases by just half a percentage point more than we would expect given a minimal set of controls (none of which are at the level of the *voting booth*). The cumulative contribution of this unexplained time trend to the overall margin is

---

[15] (Though in 2016 the OAS did not conduct an audit, as it did in 2019.)

[16] In App. F, we also show that the 2016 data create a time trend nearly identical to that of 2019.

approximately 1.1 percentage points over the entire reporting window, about 10% of Morales's overall final margin of 10.57 points. We can think of this as an upper bound on the contribution of unobserved factors to the difference between the candidates; again, these unobserved factors include (a) all voting-booth-level characteristics (b) all precinct-level characteristics except 2016 vote margin, among many others. This finding suggests that the changing composition of voting booths—rather than fraud—explains the pro-Morales shift in vote share over the reporting window.

4 How late-counted votes were interpreted as evidence of fraud, and an alternative interpretation

Despite this common-sense explanation for the vote-share trend in Bolivia's presidential election, the OAS interpreted late-counted votes as indicative of fraud.

First, using time stamps from the preliminary results system (TREP), the OAS claimed that Morales's vote share jumped discontinuously after 95% of the vote had been counted (OAS, 2019*a*, p. 88).[17] We find that this apparent jump was the artifact of an inappropriate estimator; when we follow best practices (Calonico, Cattaneo and Titiunik, 2014), Morales's vote share is continuous at 95% of the preliminary count.

Second, using time stamps from the definitive results system (*Cómputo*), the OAS claimed that Morales's vote share showed "a striking upward trend in the final 5%" (OAS, 2019*a*, p. 92). But in fact, the OAS sorted these time stamps alphabetically rather than chronologically (so that 7:01 p.m. comes right after 7:01 a.m.); when we correct this error, the "striking upward trend" disappears (as Rosnick, 2020*b*, noted in *Jacobin* magazine).

---

[17] The preliminary results system data actually contain five different time stamps corresponding to different stages of the process. In our own analysis of the pro-Morales shift in vote share above, we use the *transmission* time stamp because it reflects *reporting* time; here, we follow the OAS in using the *verification* time stamps. See Appendix A for details.

4.1 The "inexplicable" discontinuous jump in vote share: artifact of an error

It is not obvious to us whether—or under what conditions—a discontinuous jump in vote share would constitute evidence of fraud.[18] On one hand, we can easily generate innocuous explanations for such jumps; for example, if all of Philadelphia were to submit results at the same moment, the trend in Democratic vote share in Pennsylvania would undoubtedly be discontinuous at that point. On the other hand, it is at least as easy to construct theories of fraud that would produce jumps in the vote-share trend. Neither the OAS audit team (2019*a*) nor Nooruddin (2020*c*) explicitly articulated a theory of fraud that would produce the alleged jump in Morales's vote share at 95% of the vote counted, but the implicit notion was one of centralized tampering: realizing that Morales was not on track to win by more than 10 points, his agents crudely added votes in all booths that had yet to report. Hence his victory was "only made possible by a massive and unexplainable surge in the final 5% of the vote count. Without that surge ... he would not have crossed the 10% margin that is the threshold for outright victory" (OAS, 2019*a*, p. 94).

In support of this claim, the OAS presented Figure 3a.[19] But the apparent discontinuous jump in this figure—at 95% of the preliminary count—is the artifact of using an estimator inappropriate for regression discontinuity analysis. The OAS created the smoothed line in Figure 3a by estimating one local constant regression at each data point and connecting the predicted values.[20] One problem with this approach is that local constant regression often misrepresents the data at boundary points (that is, at the edges). This "boundary bias" problem is well documented: "a polynomial of order zero—a constant fit—has undesirable theoretical properties at boundary points, which is precisely where regression discontinuity estimation must occur" (Cattaneo,

---

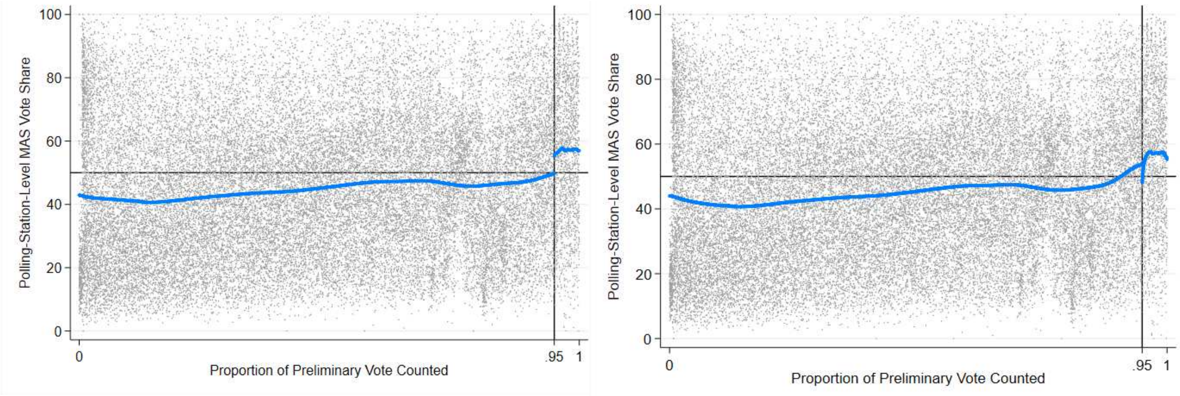
[18] Key references on election forensics do not mention discontinuous changes (e.g. Hicken and Mebane, 2017; Alvarez, Hall and Hyde, 2009).

[19] Note that the x-axis in these figures is not the time stamp itself (as in the previous section) but rather *percent of vote counted when a given voting booth's numbers were verified in the preliminary results system*. This transformation of the underlying time variable conveys an important advantage: while the time stamps themselves have long tails, the percent of vote counted is distributed nearly uniformly between zero and one. The graphs therefore visualize how vote shares change as the overall preliminary results tally progressed—not as time itself progressed. However, this transformation also entails a drawback: it artificially distorts the time trend. See Appendix Figure H.6.

[20] In particular, the OAS used Stata's `lowess` function, with the `mean` option, which implements local constant regression rather than local linear regression ("running-mean smoothing" rather than "running-line least-squares smoothing," which is the default).

Figure 3: An analytic error and the supposed jump at 95%

Figure (a) reproduces OAS (2019*a*) (p. 88). Figure (b) shows that the apparent jump disappears when we simply use local linear rather than local constant regression.



(a) Local Constant Regression
(Replication of OAS, 2019*a*, p. 88)

(b) Local Linear Regression
(No other changes)

The gray dots mark the underlying raw data. The lines mark lowess estimates with handpicked bandwidths, as implemented by Nooruddin (2020*b*). Both figures exclude the 4.4% of observations without time stamps in the preliminary results system; see Appendix B for additional discussion.

Idrobo and Titiunik, 2019, p. 38).[21] In Figure 3b, we instead use a local polynomial of degree one (i.e., local linear regression); this change alone is sufficient to eliminate the appearance of a jump.

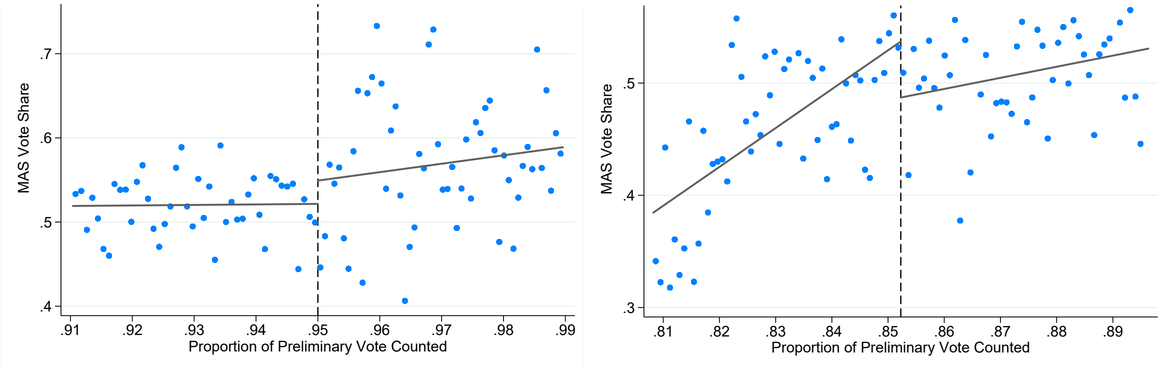
The use of local constant rather than local linear regression is not the only problem with Figure 3a. For one thing, this key figure excludes the 4.4% of observations that never made it in to the preliminary count—contrary to the OAS's claim that "all analysis conducted below includes these additional [observations]" (OAS, 2019*a*, p. 86). When we append these observations to the end of the preliminary results data, as the OAS claimed to do (p. 86), there is no discontinuity at 95% (see Appendix B). For another, the local regressions underlying Figure 3a use handpicked, arbitrary bandwidths. Moreover, the OAS presented no formal test of the null hypothesis of continuity at 95% of the preliminary count. Our simple modification in Figure 3b does not solve these problems (it merely illustrates the severe boundary bias problem created by the use of local constant regression in Figure 3a).

---

[21] See also Yu and Jones (1997), who conclude, "Detrimental boundary influence indeed exists when using local constant fitting in some cases, and it is this aspect which clinches the argument in favour of local linear smoothing" (p. 165); as well as Fan and Gijbels (1996), Sections 2.2.3, 3.2.5, and 3.4.2, and Imbens and Kalyanaraman (2011), p. 935.

Figure 4: The Absence of Discontinuities at Two Points

The points mark means of MAS's vote share in bins of 0.1 points (one tenth of one percent); the lines mark local polynomial fits with a triangular kernel and the bandwidth proposed by Calonico, Cattaneo and Titiunik (2014).



(a) At the OAS cutoff (95% of the preliminary count)

(b) At the shutdown (10/20, 7:40 p.m.)

To estimate the size of the treatment effect at 95% of the preliminary count, we use the data-driven regression discontinuity estimator proposed by Calonico, Cattaneo and Titiunik (2014). This approach estimates the treatment effect by running two local linear regressions precisely at the cutoff (one to the left, one to the right). We use this estimator to test for discontinuities at two points: (1) 95% of the preliminary count, i.e., the point studied by the OAS in Figure 3a; and (b) 7:40 p.m. on election night, when the government stopped publishing updated results (see Context section). We cannot reject the null of continuity at either of these two points.

Figure 4 graphs MAS's vote share at these two moments. The dots mark average MAS vote share in 0.1-point optimal bins; the lines plot the estimated local polynomials of degree one, with optimal bandwidth and a triangular kernel (for a comprehensive discussion, see Cattaneo, Idrobo and Titiunik, 2019). Neither presents visual evidence of a treatment effect. Table 2 reports our estimates of these treatment effects (Calonico, Cattaneo and Titiunik, 2014). Both are statistically indistinguishable from zero: we cannot reject the null of continuity at either of the cutoffs. In Appendix G, we show that this result is robust to (a) the (random) sort order within identical time stamps, (b) various choices of polynomial degree, and (c) bandwidth. Which is all to say that we find no evidence of the alleged discontinuous jump in Morales's vote share at 95% of the vote counted—the "surge" to which the OAS attributed his first-round victory (OAS, 2019*a*, p. 94).

Table 2: Non-Parametric Regression Discontinuity Estimates

Estimates of discontinuities at three points (Calonico, Cattaneo and Titiunik, 2014).

					Robust		Observations	
Cutoff	Date & Time	Sample*	RD Estimate	BW	p-val	95% C.I.	Left	Right
0.950	10/20/2019 20:03:59	Truncated	0.024	0.040	0.816	[-0.052, 0.065]	1,267	1,316
0.852	10/20/2019 19:40:57	Full	-0.028	0.044	0.374	[-0.077, 0.029]	1,455	1,457

* *Truncated* refers to the sample used by the OAS, which excludes the voting booths without time stamps in the preliminary results system. This is thus the threshold analyzed by OAS (2019*a*).

4.2 The "striking upward trend:" artifact of a coding error

Both the final audit report of the OAS (2019*a*) and follow-up commentary by the OAS consultant (Nooruddin, 2020*c*) focused on the analysis using time stamps from the preliminary results system (i.e., the analysis discussed in the previous section). This makes sense, because the preliminary-system time stamps reflect reporting time (see Appendix A), while the definitive-system time stamps reflect geographic clusters (that is, the definitive results system counts one region at a time). But the audit report nevertheless also presented results using definitive-system time stamps (*cómputo*), apparently for robustness ("we should analyze if the same patterns emerge if we use only the *cómputo* time stamps," p. 91).

The OAS claimed that "similar patterns emerge" (p. 91) in analysis using time stamps from the definitive results system. In support of this statement, the OAS presented Figure 5a, in which there is a "striking upward trend" in MAS's vote share after 95% of votes are counted in the definitive results system. But this pattern is the artifact of a coding error. The OAS sorted the definitive-system time stamps alphabetically, such that 7:01 p.m. comes right after 7:01 a.m., rather than chronologically (as Rosnick, 2020*b*, noted in *Jacobin* magazine). Correcting this error eliminates the appearance of an anomalous late-breaking surge in MAS vote share (Figure 5b).[22]

4.3 Within-precinct variation: Fraud or secular trend?

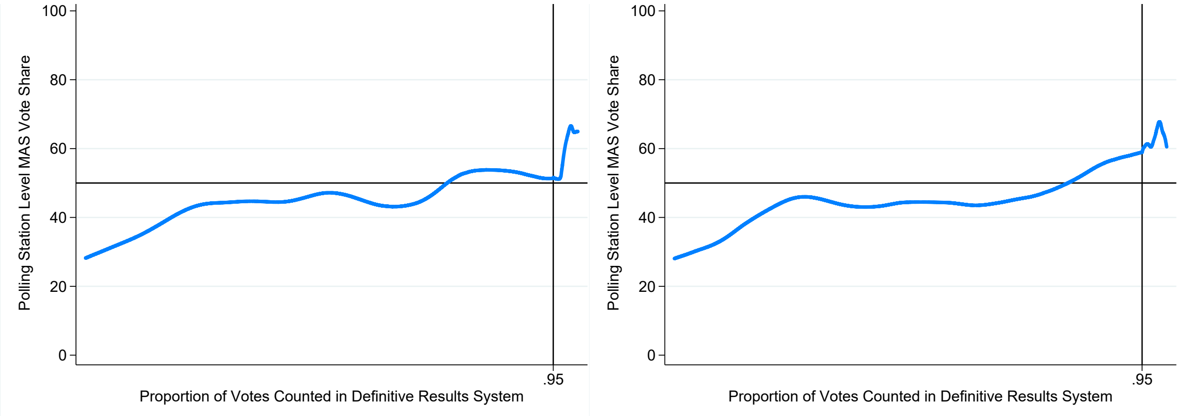
Researchers outside the OAS also pointed to late-counted votes as indicative of fraud in the Bolivian presidential election. Escobari and Hoover (2019), for example, high-

---

[22] In an earlier version of the present paper, written without access to the OAS replication materials, we did not discuss this set of results because we could not replicate them. It did not occur to us that the time stamps might have been sorted alphabetically. This error only came to light after the OAS replication materials were posted in response to our paper (Nooruddin, 2020*a*,*b*).

Figure 5: A coding error and the supposed "striking trend" (Rosnick, 2020*b*)

Figure (a) reproduces OAS (2019*a*) (p. 92), for which time stamps were mistakenly sorted alphabetically (7:01 p.m. follows 7:01 a.m.). Figure (b) shows that the apparent "striking upward trend" disappears when time stamps are sorted chronologically, as noted in the press (Rosnick, 2020*b*).


(a) Time Stamps Sorted Alphabetically
(Replication of OAS, 2019*a*, p. 92)

(b) Time Stamps Sorted Chronologically
(No other changes)

The lines mark lowess estimates with handpicked bandwidths, as implemented by Nooruddin (2020b).

light within-precinct variation. Specifically, they note that MAS performed better in voting booths reporting after the government stopped publishing updated results (*post-shutdown*) than in voting booths *from the same precinct* that reported earlier (*pre-shutdown*). Escobari and Hoover view the within-precinct variation as evidence of "a statistically significant case of electoral fraud" (p. 1); Newman, similarly, interprets it as evidence that "the OAS findings were correct" (p. 1).

In our view, these inferences are unjustified. The analysis in Escobari and Hoover (2019) and Newman (2020) compares two periods (*pre* and *post*) without accounting for a secular trend. We show in Appendix E that the within-precinct increase in MAS's vote margin begins early on election night, well before the 7:40 p.m. suspension of the publication of electoral results (note that approximately 85% of the vote had been counted at this point; it is earlier than the 95% marker studied in the previous section). Accounting for this secular trend eliminates the appearance of an anomalous within-precinct pre-post difference in vote shares.

Our analysis raises a question: why is there a secular within-precinct trend in MAS's vote margin? Why do voting booths counted later favor MAS more than voting booths *from the same precinct* that were counted earlier? In Appendix E, we propose a possible explanation. Voters' socio-economic status is unlikely to be exactly identical across voting booths within a precinct—especially because voters are assigned to booths alphabetically, not randomly, and because surnames are tied to ethnicity. Booths with voters of lower socio-economic status are more likely to vote MAS (Madrid, 2012, p. 69–72). It is also easy to understand why these booths cound report later: less-educated voters might take more time to vote, count votes, and fill out tally sheets (recall that the voting-booth *jurors* who tally votes are chosen from among each booth's registered voters). It would therefore be unsurprising to find a positive within-precinct correlation between MAS margin and time. Of course, this is not the only explanation for the secular within-precinct trend, as we discuss in Appendix E. Rather, it provides an example of one plausible explanation that does not require centralized tampering with the tally in the "post" period studied by Escobari and Hoover (2019) and Newman (2020).

Neither the secular trend nor our proposed explanations establish the absence of tampering with late-reporting booths in 2019; rather, they imply that we do not *need* electoral manipulation in order to explain the within-precinct differences that these authors cited as evidence of foul play.

5 Conclusion

The Organization of American States, Bolivian politicians, and academic researchers pointed to late-counted votes as indicative of fraud in the Bolivian presidential election of October, 2019—with dramatic political consequences (Crisis Group, 2020). We find instead that we can explain the pro-incumbent shift in vote share without invoking fraud. Most of the shift stems from just two observable characteristics of electoral precincts: region and rurality. And what looked like a late-breaking surge in the incumbent's vote share—which the OAS (2019*a*) presented as indicative of electoral malpractice—was in fact the artifact of methodological and coding errors.

Our analysis does not establish the absence of fraud in this election; that could never be determined on the basis of quantitative analysis alone. Rather, we find that the pro-incumbent shift in late-counted votes is not itself indicative of fraud. This echoes findings from the U.S. case, where researchers have similarly identified innocuous explanations for contentious late-counted votes (Foley, 2013; Foley and Stewart, 2020; Li, Hyun and Alvarez, 2020).

Our findings also speak to a general problem in election administration. Governments rarely announce election results all at once; instead, they release partial tallies as they trickle in, telling the public how things stand with 30% of precincts reporting, 70%, 90%, and so on. These updates create transparency and respond to the public's demand for information. But they also entail an important and seldom-studied cost: raising false hope. This is dangerous, because dashed hopes can spark conflict.

Incremental reporting of results thus creates a tradeoff between transparency and certainty. Diagnosing shifts in late-counted votes can lower the costs of transparency. Researchers have largely done this work for the United States (Foley, 2013; Foley and Stewart, 2020; Li, Hyun and Alvarez, 2020), but, to the best of our knowledge, we are the first to do so elsewhere. In Brazil, for example, the left candidate in the 2018 presidential election earned just 25% of votes counted early but more than 40% of votes counted late. In the Colombian presidential election that same year, Gustavo Petro fared far better as election night progressed. Do these trends stem from regional variation in the order in which votes are counted? Or from changes in the mix of urban and rural ballots? Distinguishing and publicizing these mechanisms can help protect the legitimacy of the electoral process.

Our findings suggest opportunities for future work. First, future studies could investigate the conditions under which electoral observers use quantitative analysis to study electoral integrity; as we note, the quantitative indicators applied to the Bolivian case would have revealed similar patterns in Brazil, or in the previous poll in Bolivia, both of which were endorsed by OAS missions. Second, voting technology in many countries is better suited to documenting shifts in late-counted votes than voting technology in the United States; comparative evidence on the magnitude of these shifts would provide important perspective on the Bolivian and U.S. cases. Finally, comparative work could assess which (if any) characteristics of shifts in late-counted votes *should* be interpreted as evidence of possible fraud.

References

AFP. 2019. "Morales opponent rejects poll results in Bolivia." Agence France Presse.

Aguilar, Wilson. 2018. "Renuncia de Katia Uriona deja más sombras en el TSE y las primarias." Los Tiempos.

Alvarez, Michael R., Thad E. Hall and Morgan H. Llewellyn. 2008. "Are Americans Confident Their Ballots Are Counted?" *Journal of Politics* 70(3):754–766.

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.

ANF. 2019. "Elecciones 2019: 14 mil encuestados perfilan apretada victoria de Mesa sobre Evo en balotaje." Agencia de Noticias Fides.

Animal Político. August 8, 2017*a*. "Cómputo final da triunfo a Del Mazo; Morena y el PAN van a impugnar la elección." Animal Político.

Animal Político. June 4, 2017*b*. "Conteo oficial le da ventaja a del Mazo en el Edomex; AMLO dice que ganó Morena." Animal Político.

Animal Político. June 5, 2017*c*. "Ecatepec, Tlalnepantla o Tultitlán, ¿en qué zonas del Estado de México resultó ganador Morena?" Animal Político.

Anria, Santiago. 2018. *When Movements Become Parties: The Bolivian MAS in Comparative Perspective*. Cambridge Studies in Comparative Politics Cambridge University Press.

Anria, Santiago and Jennifer Cyr. 2019. "Is Bolivia's democracy in danger? Here's what's behind the disputed presidential election." Washington Post.

Beaulieu, Emily and Susan D. Hyde. 2009. "In the Shadow of Democracy Promotion: Strategic Manipulation, International Observers, and Election Boycotts." *Comparative Political Studies* 42(3):392–415.

Birch, Sarah. 2010. "Perceptions of Electoral Fairness and Voter Turnout." *Comparative Political Studies* 43(12):1601–1622.

Bush, Sarah Sunn and Lauren Prather. 2017. "The Promise and Limits of Election Observers in Building Election Credibility." *The Journal of Politics* 79(3):921–935.

Bush, Sarah Sunn and Lauren Prather. 2018. "Who's There? Election Observer Identity and the Local Credibility of Elections." *International Organization* 72:659–92.

Calonico, Sebastian, Matias D. Cattaneo and Max H. Farrell. 2018. "On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference." *Journal of the American Statistical Association* 113(522):767–779.

Calonico, Sebastian, Matias D. Cattaneo and Rocio Titiunik. 2014. "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica* 82(6):2295–2326.

Cambara Ferrufino, Pablo Cesar. 2019. "Costas atribuye a una 'impericia' la paralización de la transmisión de votos del TREP." El Deber.

Carothers, Thomas. 1997. *The Rise of Election Monitoring: The Observers Observed.* Journal of Democracy.

Cattaneo, Matias D., Nicolás Idrobo and Rocío Titiunik. 2019. *A Practical Introduction to Regression Discontinuity Designs.* Cambridge University Press.

Cattaneo, Matias D., Richard K. Crump, Max H. Farrell and Yingjie Feng. 2019. "On Binscatter." *arXiv* .

Collyns, Dan. 2019. "Evo Morales agrees to new elections after irregularities found." The Guardian.

Comunidad Ciudadana. November 8, 2019. "Carta abierta al Vicepresidente Álvaro García Linera." Twitter.

Corrales, Javier. 2016. "Can Anyone Stop the President? Power Asymmetries and Term Limits in Latin America, 1984–2016." *Latin American Politics and Society* .

Cottrell, David, Michael Herron and Sean Westwood. 2018. "An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election." *Electoral Studies* 51:123–142.

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

CSC. November 4, 2019. "Carta al Presidente Evo Morales." Comité pro Santa Cruz.

Donno, Daniela. 2010. "Who Is Punished? Regional Intergovernmental Organizations and the Enforcement of Democratic Norms." *International Organization* 64:593–625.

Donno, Daniela. 2013. *Defending Democratic Norms: International Actors and the Politics of Misconduct*. Oxford University Press.

Escobari, Diego and Gary Hoover. 2019. "Evo Morales and Electoral Fraud in Bolivia: A Natural Experiment Estimate." Working Paper.

Ethical Hacking. 2019. "Informe Consolidado de Producto 1 y Producto 2: Línea de Tiempo de la Consultoría al OEP." Ethical Hacking.

Exeni Rodríguez, José Luis. 2020. "Interview with José Luis Exeni Rodríguez." *Former Vice President of the Tribunal Supremo Electoral* .

Falleti, Tulia and Emilio Parrado. 2018. *Latin America Since the Left Turn*. Univeristy of Pennsylvania Press.

Fan, J. and I. Gijbels. 1996. *Local Polynomial Modelling and its Applications*. CRC Press.

Flores, Paola and Carlos Valdez. 2019. "Bolivia says OAS team to audit disputed election." AP News.

Foley, Edward B. 2013. "A big blue shift: Measuring an asymmetrically increasing margin of litigation." *Journal of Law and Politics* .

Foley, Edward B. and Charles Stewart. 2020. "Explaining the Blue Shift in Election Canvassing." Working Paper.

Forebears.io. 2020. "Most Common Last Names in Bolivia." Forebears 2012–2020.

Fujiwara, Thomas. 2015. "Voting technology, political responsiveness, and infant health: evidence from Brazil." *Econometrica* 83(2):423–464.

Goel, Sharad, Marc Meredith, Michael Morse, David Rothschild and Houshmand Shirani-Mehr. 2020. "One Person, One Vote: Estimating the Prevalence of Double Voting in U.S. Presidential Elections." *American Political Science Review* pp. 1–14.

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

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

Hyde, Susan D. 2007. "The Observer Effect in International Politics: Evidence from a Natural Experiment." *World Politics* 60(1):37–63.

Hyde, Susan D. 2011. *The Pseudo-Democrat's Dilemma: Why Election Observation Became an International Norm.* Cornell University Press.

Hyde, Susan D. and Nikolay Marinov. 2014. "Information and Self-Enforcing Democracy: The Role of International Election Observation." *International Organization* 68(2):329–359.

Imbens, Guido and Karthik Kalyanaraman. 2011. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *The Review of Economic Studies* 79(3):933–959.

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

Kanyinga, Karuti. 2009. "The legacy of the white highlands: Land rights, ethnicity and the post-2007 election violence in Kenya." *Journal of Contemporary African Studies* 27(3):325–344.

Kavakli, Kerim Can and Patrick M. Kuhn. 2020. "Dangerous Contenders: Election Monitors, Islamic Opposition Parties, and Terrorism." *International Organization* 74(1):145–164.

Kelley, Judith. 2009. "D-Minus Elections: The Politics and Norms of International Election Observation." *International Organization* 63:765–87.

Kelley, Judith. 2012. *Monitoring Democracy: When International Election Observation Works, and Why it Often Fails.* Princeton University Press.

Klein, Herbert S. 2011. *A Concise History of Bolivia*. Cambridge Concise Histories 2 ed. Cambridge University Press.

Kurmanaev, Anatoly, Mónica Machicao and Ernesto Londoño. 2019. “Military Calls on President to Step Down After Election Dispute in Bolivia.” The New York Times.

Kurmaneav, Anatoly and Cesar Del Castillo. 2019. “How and Unknown Female Senator Came to Replace the Bolivian President Evo Morales.” The New York Times.

Li, Yimeng, Michelle Hyun and R. Michael Alvarez. 2020. “Why Do Election Results Change After Election Day? The ”Blue Shift” in California Elections.” *American Government and Politics* Forthcoming.

Los Tiempos. 2019*a*. “El sistema de Transmisión de Resultados Electorales Preliminares (TREP) se actualizó al 94,99una diferencia de 10 puntos entre los candidatos del MAS y CC.” Twitter.

Los Tiempos. 2019*b*. “TSE dice que se paralizó el TREP para ”evitar la confusión” con resultados de departamentos.” Los Tiempos.

Madrid, Raúl. 2012. *The Rise of Ethnic Politics in Latin America*. Cambridge University Press.

Mebane, Walter. 2019. “Evidence Against Fraudulent Votes Being Decisive in the Bolivia 2019 Election.” Working Paper.

Mesa, Carlos D. 2019. “Post-election thread.” Twitter.

Minoldo, Sol and Rodrigo Quiroga. 2020. “¿Hubo fraude en las elecciones de Bolivia?” Cenital.

NEOTEC. October 28, 2019. “Informe Final Elecciones Generales 2019.” NEOTEC.

Newman, John. 2020. “The OAS Conclusions about the election integrity of the Bolivian election are correct.” Working Paper.

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.

Norris, Pippa. 2014. *Why Electoral Integrity Matters*. Cambridge University Press.

Norris, Pippa, Holly Ann Garnett and Max Grömping. 2020. "The paranoid style of American elections: explaining perceptions of electoral integrity in an age of populism." *Journal of Elections, Public Opinion and Parties* 30(1):105–125.

OAS. 2016*a*. "Informe al Consejo Permanente, Misión de Observación Electoral, Referendo Constitucional." *Organization of American States* .

OAS. 2016*b*. "Informe Preliminar: Misión OEA felicita la alta participación en el referendo constitucional de Bolivia." *Organization of American States* .

OAS. 2018. "OAS Electoral Observation Mission Congratulates Brazil for Holding its General Elections and Announces that It Will Observe the Second Round." *Organization of American States* .

OAS. 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* .

OAS. October 21, 2019*d*. "Statement of the OAS Electoral Observation Mission in Bolivia." *Organization of American States* .

OAS. October 23, 2019*c*. "Informe Preliminar de la Misión de Observación Electoral en Bolivia." *Organization of American States* .

OEP. 2019. "Cartilla de información pública: ¿Cómo funcionan los sistemas del TREP y del cómputo oficial?" *Órgano Electoral Plurinacional* .

Robinson, P. M. 1988. "Root-N-Consistent Semiparametric Regression." *Econometrica* 56(4):931–954.

Rodriguez-Larralde, Alvaro, José Dipierri, Emma Alfaro Gomez, Chiara Scapoli, Elisabetta Mamolini, Germano Salvatorelli, Sonia De Lorenzi, Alberto Carrieri and Italo Barrai. 2011. "Surnames in Bolivia: A study of the population of Bolivia through isonymy." *American Journal of Physical Anthropology* 144(2):177–184.

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* .

Simpser, Alberto. 2012. "Does Electoral Manipulation Discourage Voter Turnout? Evidence from Mexico." *The Journal of Politics* 74(3):782–795.

Simpser, Alberto and Daniela Donno. 2012. "Can International Election Monitoring Harm Governance?" *The Journal of Politics* 74(2):501–513.

TSE. 2019. "Verificación del llenado del acta." Tribunal Supremo Electoral de Bolivia Video.

UNICEF. 2014. "El derecho a la educación y a la protección en Bolivia." UNICEF.

Williams, Jack and John Curiel. 2020. "Analysis of the Quick Count in the 2019 Bolivia Election." Working Paper.

Yu, Keming and M.C. Jones. 1997. "A comparison of local constant and local linear regression quantile estimators." *Computational Statistics & Data Analysis* 25(2):159 – 166.

Appendix

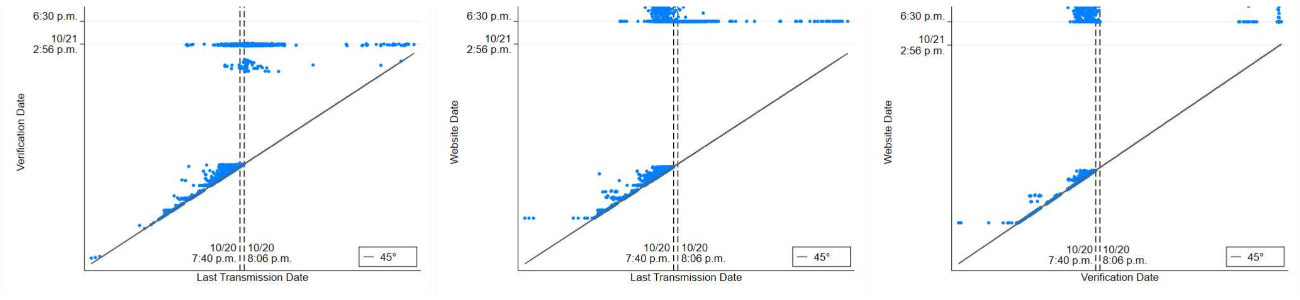
A Data: Details on time stamps

The preliminary results system (TREP) data contain five different time stamp variables: first registration, last registration, first transmission, last transmission, and verification.[23] (See the Data section for details on the preliminary results system.)

In Section 4 of the main text, we use *percentiles of verification date* as the time variable. We do this because it allows us to most closely replicate the figures in OAS (2019*a*); none of the other transformed time stamps generate graphs that look like those of the OAS. But *verification date* has two related drawbacks. First, it is itself discontinuous: the verification time stamps stop at 8:06 p.m. on election night and pick up again the following morning at 10:37 a.m. (October 21). Second, while the penultimate time stamp—last transmission date—is highly predictive of verification time prior to 8:06 p.m. on election night, this correlation breaks down after 8:06 p.m. That is, the tally sheets transmitted at 8:07 p.m., 8:08 p.m., etc., were not necessarily the first tally sheets verified on the morning of October 21.

Figure A.1: Website Times and Verification Times vs. Reporting Time

Figure (a) plots the *verification* time stamp—the one used in the main text—against the *last transmission* time stamp. The latter is continuous, whereas verification stops at 8:06 p.m. on election night and continues at 10:37 a.m. the next morning. Figure (b) plots the public (website) time stamp against the last transmission time stamp, revealing that all tally sheets transmitted after 7:40 p.m.—and many transmitted before—were published online at 6:30 p.m. the next day. Figure (c) plots the website time stamp against the verification time stamp, again revealing that many tally sheets *verified* soon after the polls closed were published online on the evening of the next day.


(a) Verification vs. Transmission (b) Website vs. Transmission (c) Website vs. Verification

All figures exclude one outlier, an observation that was not *verified* until October 22.

There is a clear reason for this. The preliminary results system was set up such that the *verification* operators would receive the transmitted tally sheet images in

---

[23] There is also a sixth, *approval*, which is missing for almost all observations.

a random order (NEOTEC, October 28, 2019, p. 5). Early in the evening, this randomization only slightly disrupted the relationship between submission time and verification time, because the pool of unverified images was fairly small. But the 7:40 p.m. shutdown of the preliminary results system halted verifications without halting transmissions. When verifications resumed the following day, tally sheet images were drawn randomly from a large pool—breaking the correlation between transmission and verification times.

Figure A.1a visualizes both of these issues, plotting the *verification* time stamp against the *last transmission* time stamp. Again, the latter is continuous while the former stops after 8:06 p.m. on election night. Moreover, while *last transmission* time strongly predicts *verification* time prior to 8:06 p.m., it does not predict verification time for those tally sheets verified the next day. Likewise, Figures A.1b and A.1c reveal that the order in which tally sheets were published online reflects the order in which they were transmitted only for a subset of early-reporting voting booths. Finally, A.2 clarifies that more than 85% of the vote was counted before the 8:06 p.m. interruption in verification.

Figure A.2: Progress of the Preliminary Results System (TREP)

This figure excludes the 4.4% of voting booths (4.1% of the vote) without TREP time stamps.

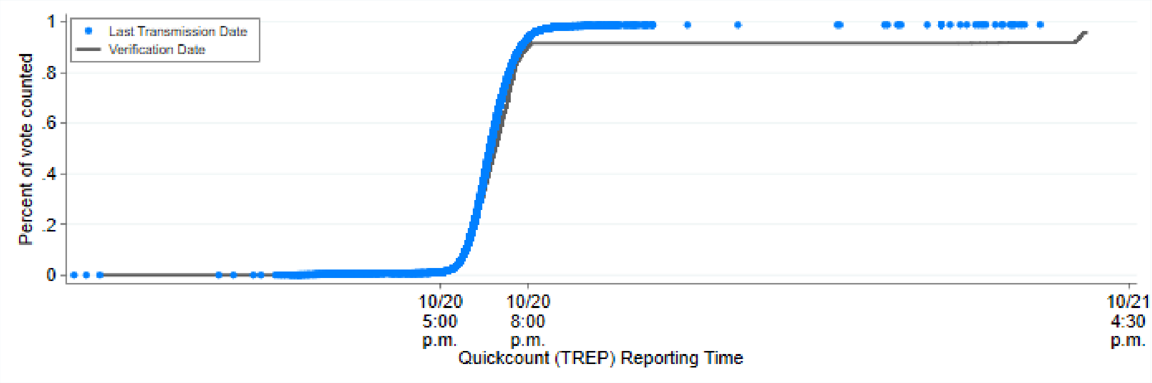
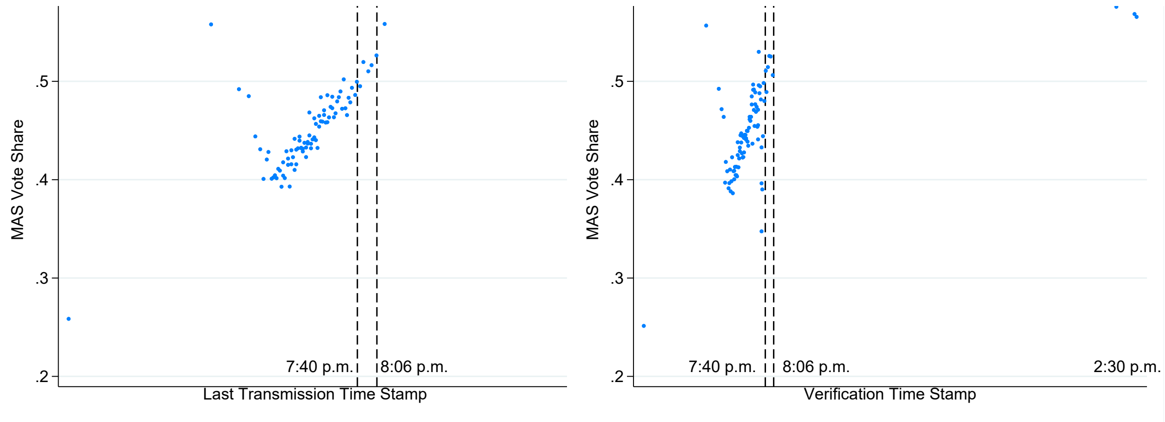


Figure A.3 shows how MAS's vote share changes as a function of clock time (rather than as a function of *percentile of vote verified*, which is what we study in the main text). The points mark average MAS vote share in optimal (data-driven) bins of the timestamp variables (Cattaneo et al., 2019). The vertical lines mark two times of interest: 7:40 p.m., when the government stopped publishing updated results; and 8:06 p.m., when the verification time stamps stop until the following morning. (As noted above, transmission time stamps continue through the night).

Figure A.3: Last Transmission Date, Verification Date, and MAS Vote Share

Figure (a) plots average MAS vote share in bins of the preliminary results system *last transmission date*, using the optimal (data-driven) bins (Cattaneo et al., 2019). Figure (b) plots average MAS vote share in optimal bins of the preliminary results system *verification date*.


(a) Last Transmission Time Stamp

(b) Verification Time Stamp

Figure A.3b highlights a potential problem with testing for discontinuities in these data. The discontinuity analysis uses *percentiles* of the verification time stamp, which is to say, a transformation of verification date that places the 8:06 p.m. time stamps right next to the 10:37 a.m. (next day) time stamps, effectively closing the long gap in the actual time series. Thus if we were to test for a discontinuity at 8:06 p.m. using percentiles of verification date, we would be testing for a discontinuity in MAS vote share at a moment when the running variable (time) is itself discontinuous. Worse, as noted above, the order in which tally sheets were verified on the morning of October 21 is only loosely related to the order in which they were transmitted the night before. As it happens, neither of the moments studied by the OAS—7:40 p.m., and 95% of the vote counted—coincide with 8:06 p.m. on election night (see Table 2), so this problem does not arise.

Using the last transmission time stamp, Figure A.3a reveals an apparently smooth trend in MAS vote share before and after 8:06 p.m. When we test for a discontinuity in MAS vote share at 8:06 p.m., using last transmission date and again following Calonico, Cattaneo and Titiunik (2014), the point estimate is positive (4.8 percentage points) but statistically indistinguishable from zero.

Overall, it strikes us that the last transmission time stamp better captures the reporting sequence, while the verification time stamp perhaps better captures the counting/tabulation sequence. Again, we focus on the verification time stamp in the main text because it allows us to replicate the OAS results. However, the substantive takeaway from our own analysis in Section 3.3 remains unchanged when we use the last transmission time stamp: MAS performs predictably well, *not* surprisingly well, in the period after the government stopped publishing updated results.

B The 4.4% of observations excluded from the preliminary results

The analysis in OAS (2019*a*) focuses on a data set that merges the preliminary results system (TREP) time stamps with the definitive-system (*cómputo*) vote tallies, at the level of the voting booth. In our view, this makes sense. The preliminary-system time stamps capture when each voting booth's tallies were verified, which is the relevant time series for investigating the shift in late-counted votes; the *cómputo* vote tallies are those that determine the final margin.

Using the preliminary-system time stamps entails a challenge: how to treat the set of voting booths that never transmitted results through the preliminary system (TREP). These 1,513 voting booths account for 4.4% of all observations, and they are excluded from the preliminary system for diverse reasons (including lack of cellular service and tally sheet illegibility). Regardless, their preliminary-system time stamps are unobserved (perhaps even undefined).

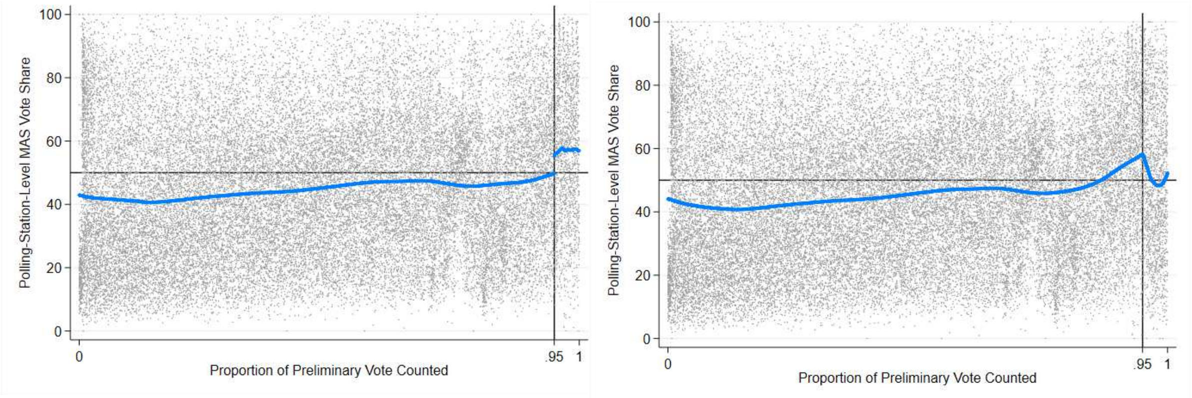
In our own analysis in Section 3, we treat these voting booths as the latest reporters within their respective municipalities; in Appendix C, we show robustness to several approaches.

The text of the OAS audit report claims to treat these voting booths as "late reporters" (p. 86), under the assumption that they finished tallying only after the preliminary results system closed. The report states: "All the analysis conducted below include these additional polling stations. Since they were not included in the TREP [preliminary system], they are treated as being late reporters" (p. 86). We interpreted this to mean that OAS (2019*a*) sorted the first 33,038 booths by their preliminary results system time stamps, and then appended the remaining 1,513 voting booths (4.4%) at the end, presumably in a random order.

In an earlier draft of this paper, we alleged that, rather than append the "late-reporting" voting booths to the end of the data set as claimed, the OAS dropped them when creating Figure 3a. This is a consequential exclusion. The "late-reporters" account for 4.4% of tally sheets and 4.1% of votes, which is to say, the vast majority (82%) of the last 5% of votes counted (if we assume, as the OAS does, that they were late reporters). Any analysis focused on the last 5% of votes counted will therefore be quite sensitive to the treatment of the booths without preliminary results system

Figure B.1: Exclusion of “late reporters” and the jump at 95%

Figure (a) reproduces OAS (2019*a*) (p. 88). Figure (b) shows that the apparent jump disappears when we append the observations without preliminary-system time stamps.



(a) Dropping “Late Reporters” (Replication of OAS, 2019*a*, p. 88)

(b) Including “Late Reporters” (No other changes)

The gray dots mark the underlying raw data. The lines mark nonparametric fits using the tricube weighting function and the bandwidths handpicked in Nooruddin (2020*b*), namely, 0.3 to the left of the cutoff and 0.6 to the right of the cutoff.

time stamps.

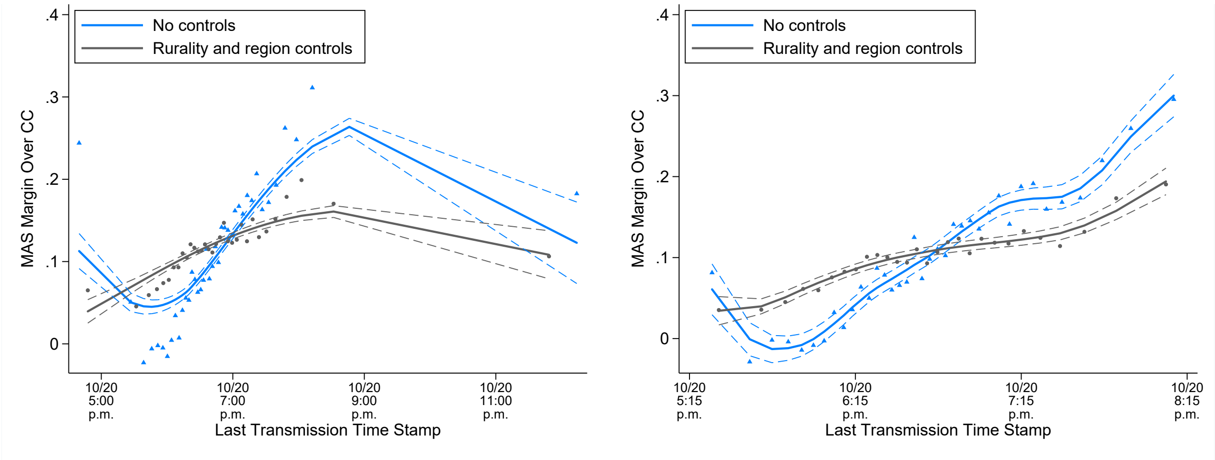
The OAS replication materials (Nooruddin, 2020*b*), posted in response to an earlier draft of the present paper, confirm that the “late-reporting” voting booths were in fact dropped in creating Figure 3a (i.e., the graph presented as evidence of a discontinuous jump in MAS’s vote share). If we include the “late reporters” at the end, as the OAS audit report claimed to do, we obtain Figure B.1b. In this case, there is neither a jump nor an uptick in the trend of MAS’s vote share in the final 5% of the count.

In his response to our earlier draft, Nooruddin (2020*a*) argued that the OAS audit report never *claimed* to include the “late-reporting” voting booths in this key results figure. We maintain that the language of the report implies otherwise. Regardless, excluding the “late-reporters” from the key results figure in an analysis of late-counted votes strikes us as unfortunate—whether by choice or by mistake.

C Additional semiparametric results

Figure C.1: Additional semiparametric results

This figure shows the robustness of our semiparametric results to (a) including all time outliers in the sample; (b) dropping the observations without time stamps in the preliminary results system; (c) treating them as early reporters in their respective municipalities; and (d) treating them as typical (median) reporters in their respective municipalities.

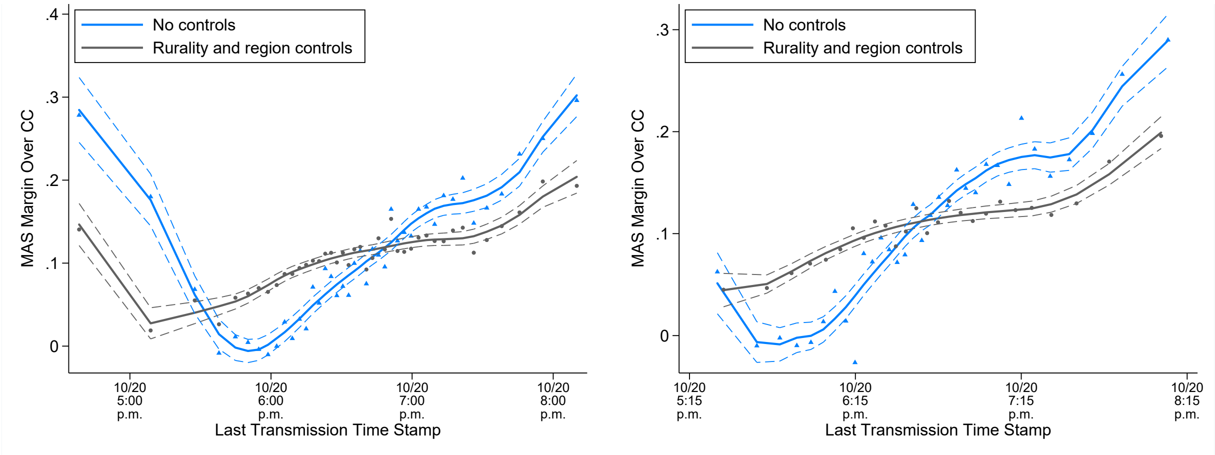


(a) No trimming

(b) Dropping missings

(c) Replacing at the minimum

(d) Replacing at the median



Points mark the average MAS margin over Civic Community in optimal (data-driven) bins of the time variable (Cattaneo et al., 2019). The solid blue fits mark a local-polynomial estimation that follows Calonico, Cattaneo and Farrell (2018). The grey lines mark an estimate of $f(\text{Time})$ from Equation 1, using the semi-parametric estimator proposed in Robinson (1988).

D Escobari and Hoover (2019) Replication

In the main text, we note a problem with the specification in Escobari and Hoover (2019): it includes an indicator for *post* without accounting for a secular (within-precinct) trend in MAS's vote margin. We show that when we account for this trend, the coefficient on *post* is estimated at zero.

The results presented in Table E.1, Column (3) in the main text—reproduced in Column (3) of Table D.1 below—do not exactly replicate Escobari and Hoover (2019). Our coefficient on *post* is estimated at 0.0056 (about half of one percentage point), whereas theirs is estimated at 0.0037. The principal difference is that Escobari and Hoover use what we call the *website* time stamps (see previous section, Appendix A), whereas we use the internal *verification* time stamps. When we use the *website* time stamps, as in Column (5) of Table D.1, we can replicate their result almost exactly.

Table D.1: Replication of Escobari and Hoover (2019)

Estimates of Eqn. 2. The D.V. is MAS's margin over Civic Community (scaled 0–1).

	Last Transmission		Verification		Website	
	(1)	(2)	(3)	(4)	(5)	(6)
Post shutdown (0/1)	0.0048 (0.0019)	-0.0018 (0.0023)	0.0056 (0.0019)	0.0001 (0.0023)	0.0038 (0.0018)	0.0036 (0.0018)
Reporting time percentile		0.0164 (0.0034)		0.0130 (0.0033)		
Observations	32,946	32,946	32,946	32,946	32,925	32,946
Precinct FEs	✓	✓	✓	✓	✓	✓

Standard errors, clustered by precinct, in parentheses. Column (5) uses MAS's margin as Escobari and Hoover (2019) calculated it; Column (6) uses MAS's margin as it appears in the final tally.

A secondary difference is that Escobari and Hoover calculate MAS's margin based on a preliminary count of valid votes (the one published on the website), whereas we calculate MAS's margin based on the final count of valid votes. Because the number of valid votes differs only for 2.75% of observations, and because these differences are quite small, this alone makes little difference for the final estimates: Column (5) of Table D.1 uses the website count of valid votes; Column (6) uses the final count of valid votes. The point estimate changes by 0.0002.

E More on the within-precinct trend

Documenting the trend. Before studying within-precinct variation in vote shares, we note that the within-precinct variation in reporting time is substantial. 70% of precincts have more than one voting booth; among these precincts, the median within-precinct standard deviation in reporting time is 35 minutes—more than one fourth of the active reporting window (see Appendix Figure A.2). Moreover, 26% of precincts—and 37% of precincts with more than one voting booth—contain booths reporting before *and* after the public information blackout.

Figure E.1a presents an example of within-precinct variation; the blue diamonds mark MAS's margin in each of the 40 voting booths in a single precinct in the town of Llallagua, Potosí. In this example, MAS's margin increases with reporting time even before the government stopped transmitting updated results (at 7:40 p.m., with 85.2% of the vote verified).[24] This is not an isolated case. Let m_{bp} denote MAS's margin in voting booth b in precinct p, and $\overline{\mathrm{m}}_p$ denote the average margin in precinct p. Then Figure E.1b reveals that the residual MAS vote margin $\mathrm{m}_{bp} - \overline{\mathrm{m}}_p$ increases with reporting time.

Critically, the within-precinct divergence between MAS and CC does not accelerate after the shutdown of the public preliminary results system. If anything, the candidates' fortunes diverge more slowly after 7:40 p.m. (This fact is robust to bandwidth choice, as we show in Appendix Figure H.5).

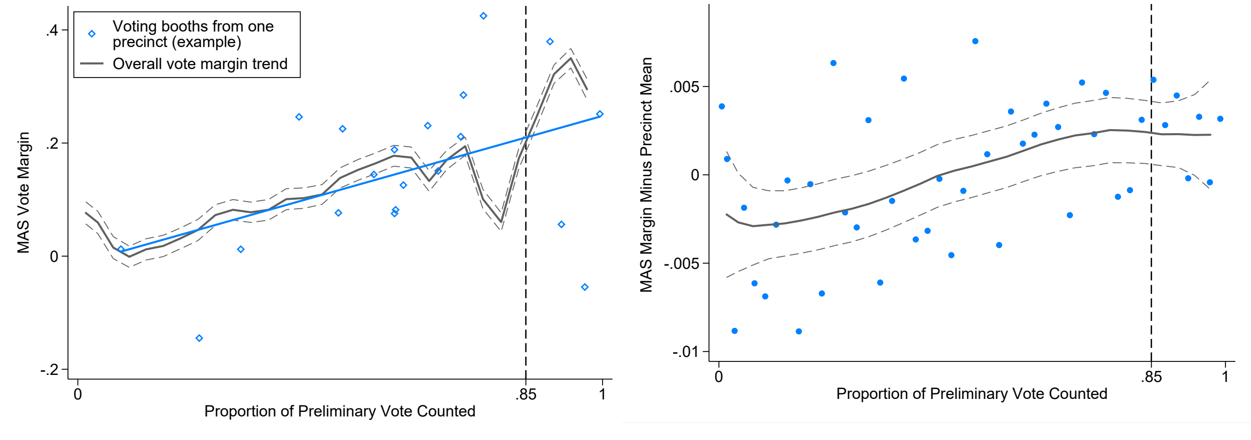
The time trend in Figure E.1b highlights a problem with the interpretation of results in Escobari and Hoover (2019). They regress MAS's margin on an indicator for *post-shutdown* and precinct fixed effects, finding that the coefficient on *post-shutdown* is positive and significant even with precinct fixed effects included. The magnitude

---

[24] The gray line in Figure E.1a, which marks the overall time trend in MAS margin, differs from the time trend in Figure 2 for two reasons. First, Figure E.1a uses the *verification* time stamp (following the work we replicate in this section), while Figure 2 uses the *transmission* time stamp, which better captures reporting time (Appendix A). The sharp non-monotonicity in Figure E.1a is caused by a server backup that produced a burst of verifications of tally sheets from the anti-Morales department of Santa Cruz. Second, Figure E.1a uses *percentiles* of reporting time on the x-axis (again following other work), while Figure 2 uses clock time. Using percentiles has the effect of visually compressing the long tails of the distribution of clock time: many more minutes elapsed between the 95th and the 96th percentiles than between the 65th and 66th percentiles; Figure E.1a obscures this fact. Indeed, Nooruddin (2020*c*) and OAS (2019*a*) commented on "the steep slope" of the trend around 7:40 p.m.; this apparent change in slope is an artifact of using *percentile of reporting time* rather than *reporting time*, as we show in Appendix Figure H.6.

Figure E.1: Within Precincts, MAS Vote Margin Increases with Reporting Time

Figure (a) provides an example of within-precinct variation; the blue diamonds mark MAS's vote margin in each of the 20 voting booths in a single precinct in the town of Llallagua. The gray line marks the overall margin trend; it differs from the trend in Figure 2 for two reasons, explained in Footnote 24. Figure (b) plots the voting-booth-level MAS margin after subtracting the precinct mean (i.e., $\mathrm{m}_{bp} - \overline{\mathrm{m}}_p$).



(a) Within-Precinct Variation: An Example

(b) Residual MAS Margin ($\mathrm{m}_{bp} - \overline{\mathrm{m}}_p$)

Local linear fits with the rule-of-thumb bandwidth from Calonico, Cattaneo and Farrell (2018)

of the coefficient is consistent with our Figure E.1b; it reveals that MAS's post-shutdown vote margin was approximately four tenths of a percentage point larger than MAS's pre-shutdown margin in the same precincts. But the Escobari and Hoover (2019) specification does not account for the secular trend in Figure E.1b: even within precinct, voting booths that report later favor MAS, even before the shutdown. Adding a time trend to the regression in Escobari and Hoover (2019) reduces the estimate of the post-shutdown increase to zero.

To see this, consider a regression of the form:

$$\begin{aligned} \mathrm{M}_{bp} = \gamma_p + \beta_1(\text{Time percentile})_{bp} + \beta_2 \mathbb{1}(\text{Post shutdown})_{bp} \\ + \beta_3(\text{Percentile} \times \text{Post})_{bp} + \epsilon_{bp} \end{aligned} \tag{2}$$

where M_{bp} is MAS's margin over CC in voting booth b in precinct p; γ_p are precinct fixed effects; $(\text{Time percentile})_{bp}$ is the percent of the vote counted when voting booth b was verified in the preliminary results system (TREP); $(\text{Post shutdown})_{bp}$ takes a value of 1 if voting booth b reported after the government stopped publishing updated results (7:40 p.m.) and 0 otherwise; $(\text{Percentile} \times \text{Post})_{bp}$ interacts $(\text{Time percentile})_{bp}$ with $(\text{Post shutdown})_{bp}$; and ϵ_{bp} is a voting-booth-specific error term.

Table E.1: Within Precinct, MAS Margin Does Not Grow Faster Post-Shutdown

Estimates of Equation 2. The dependent variable is MAS's margin over Civic Community (scaled -1 to 1). Column (1) reveals that the (linearized) growth in MAS's margin does accelerate after the shutdown; Column (2) shows that this is not true of within-precinct variation; Column (3) replicates Escobari and Hoover (2019, Table 3, Col. 3), showing that omitting the within-precinct secular trend in MAS margin produces a positive and (marginally) significant coefficient on the post-shutdown dummy; and Column (4) adds the time trend, revealing that, in this specification, the coefficient on *post-shutdown* is estimated at zero.

	(1) No Precinct FEs	(2) + Precinct FEs	(3) No time trend[§]	(4) + time trend
$\hat{\beta}_1$: Reporting time percentile[†]	0.173	0.014		0.013
	(0.02)	(0.003)		(0.003)
$\hat{\beta}_2$: Post shutdown (0/1)	0.102	0.004	0.006	0.000
	(0.02)	(0.003)	(0.002)	(0.002)
$\hat{\beta}_3$: Percentile × Post	-0.019	-0.052		
	(0.2)	(0.04)		
Observations	34,551	32,946	32,946	32,946
Precinct FEs		✓	✓	✓

Standard errors, clustered by precinct, in parentheses. [§]This is the specification in Escobari and Hoover (2019); see Appendix D for discussion. [†]For ease of interpretation of the coefficients, we center the reporting time percentile at the moment of the shutdown (7:40 p.m. on election night). Thus the coefficient on *reporting time percentile* can be interpreted as the slope of MAS's vote share before the shutdown, the coefficient on *Post* is the estimated jump (new intercept) after the shutdown, and the coefficient on the interaction term is the increase in slope after the shutdown.

Column (1) of Table E.1 reports estimates of a version of Equation 2 that *excludes* precinct fixed effects; in this specification, MAS's margin grows faster after the government stopped publishing updated results. But when we include precinct fixed effects, in Column (2), MAS's margin grows no faster after than before the shutdown. If anything, and again consistent with Figure E.1b, the growth in MAS's margin slows after the shutdown ($\widehat{\beta}_3$ is negative but imprecisely estimated).

Column (2) of Table E.1 also reveals that, *even within precinct*, there is a secular increase in MAS's margin over the reporting window. This is captured in the positive and significant coefficient on β_1. And this is the problem with the conclusions Escobari and Hoover (2019): if we omit that secular trend, as in Column (3), then the coefficient on the *post shutdown* is positive and significant.[25] When we include the secular trend,

---

[25] The estimate in Column (3) of Table E.1 is larger than the corresponding estimate in Escobari and Hoover (2019), because we use slightly different time stamps to construct the *post* variable. When we use the same time stamps, we can replicate Escobari and Hoover's estimate, as we show

as in Column (4), the coefficient on *post shutdown* is estimated at zero. The same would be true of an indicator for any artificial *post* period: post-50% of the count, post-70% of the count, et cetera. In other words, because of the within-precinct secular trend in MAS margin, the specification that Escobari and Hoover propose as a "natural experiment" is not, in fact, a natural experiment.

Possible explanations. As noted in the Context section, voting-booth jurors (*jurados*) are chosen randomly from among each voting booth's voters—not from among voters in the whole precinct. At the close of voting, the jurors count the ballots and fill out a paper tally sheet (*acta*). This aspect of electoral administration in Bolivia could easily generate a correlation between MAS vote margins and verification time. Voters' socio-economic status is unlikely to be exactly identical across voting booths within a precinct. Booths with voters of lower socio-economic status and lower levels of education are more likely to vote MAS (Madrid, 2012, p. 69–72). It is easy to imagine why those booths might also report later: voters with lower levels of education may take more time to vote; moreover, jurors with lower levels of education would likely take more time to count votes and fill out the tally sheet. It is therefore unsurprising that we find a positive within-precinct correlation between MAS margin and time.

These differences across voting booths within a precinct are likely greater because voters are assigned alphabetically—not randomly—to voting booths within precincts, as in much of the United States (Exeni Rodríguez, 2020). Of course, surname is related to ethnicity, which is related to socio-economic status in Bolivia (including education, see UNICEF, 2014, p. 30)—and indigenous surnames are distributed differently throughout the alphabet than non-indigenous surnames. Indigenous surnames are more likely to begin with C, H, or Y, for example, while non-indigenous names are more likely to begin with F, R, or S (Forebears.io, 2020). For that reason, different voting booths likely have different proportions of indigenous voters.

To illustrate, consider a hypothetical precinct with the mean number of voting booths (6.5). Each voting booth has approximately 15% of the precinct's voters. Consider two clusters of last names: those that begin with the letter C, which includes 15.9% of the population, and those that begin with R or S, which together cover 14% (Forebears.io; see also Rodriguez-Larralde et al., 2011). This hypothetical precinct

---

in Appendix D.

could then have one voting booth in which all voters' surnames begin with C, and another in which all voters' surnames begin with R or S. These booths would likely have very different proportions of indigenous voters: among the 911 most common surnames (which account for 88% of the population), 33.1% of people with C surnames have indigenous surnames, while 1.4% of the people with R or S surnames have indigenous surnames. It would therefore be unsurprising if MAS performed better in the C voting booth than in the $R + S$ voting booth; nor would it be surprising if the C voting booth reported later than the $R + S$ voting booth.

One implication of this hypothesis is that, even *within* precinct, the proportion of null ballots would be correlated with reporting time. While *blank* ballots might be interpreted as protest votes, null ballots occur when the voter makes a mistake (for example, marking two candidates instead of one). Less-educated voters are more likely to cast these ballots (Fujiwara, 2015). Thus, if within-precinct variation in voters' socioeconomic characteristics is correlated with within-precinct variation in verification time, we would also expect within-precinct variation in null ballots to be correlated with within-precinct variation in verification time. We show graphically that it is (Appendix Figure H.3).

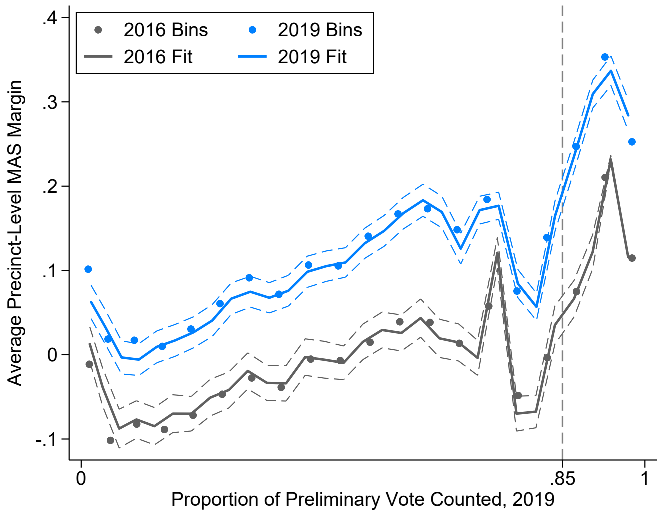
Another possible explanation for the within-precinct trends in MAS margin and in null ballots is that pro-MAS jurors strategically invalidate ballots cast for the opposition, and that doing so takes time. Writing and estimating a model to adjudicate between these explanations strikes us as a worthy objective for future work. In any case, decentralized invalidation of opposition votes throughout election night does not resemble mechanics implicitly alleged by Escobari and Hoover (2019) and Newman (2020), in which the government stopped publishing results in order to enable centralized tampering with vote tallies in late-counted voting booths.

F Comparison with 2016

It is not possible to match *voting booths* across elections, because of how the booth identifiers changed. However, we can match *precincts* across elections (Minoldo and Quiroga, 2020, show a high correlation between 2016 and 2019 precinct-level vote shares). We then calculate *average precinct-level MAS vote margin* for each voting booth in each election ($\overline{\text{m}}_p$), and plot these average precinct-level margins against each voting booth's 2019 reporting time percentile.[26] (We do not have time stamps for 2016; the figure uses 2019 time stamps for both years).

Figure F.1: 2016 Electoral Returns Reveal Similar Patterns

This figure plots average precinct-level MAS vote margins $\overline{m}_p$ in two elections against each voting booth's percentile of reporting time in 2019 (i.e., the x-axis values are the same).



We only include observations corresponding to voting booths present in both 2016 and 2019 (i.e., the samples are the same). Lines mark local linear fits using the rule-of-thumb bandwidth from Calonico, Cattaneo and Farrell (2018); the dotted lines mark 95% confidence intervals, and the bins are obtained using Cattaneo et al. (2019).

Figure F.1 plots the result: the shape of the vote-share trend appears nearly identical if we use 2016 vote margins rather than 2019 vote margins. In other words, features that the OAS flagged as anomalous in 2019 also emerge in analysis of data from 2016, an election for which the OAS congratulated Bolivia and praised the leadership of the electoral authority (OAS, 2016*a*,*b*).

---

[26] To be clear, the y-axis values in Figure F.1 are not the voting-booth-specific MAS margins m_{bp}, but rather the average precinct-level MAS margin $\overline{\text{m}}_p$. In other words, all booths b in a given precinct p have the same y-axis value in Figure F.1.

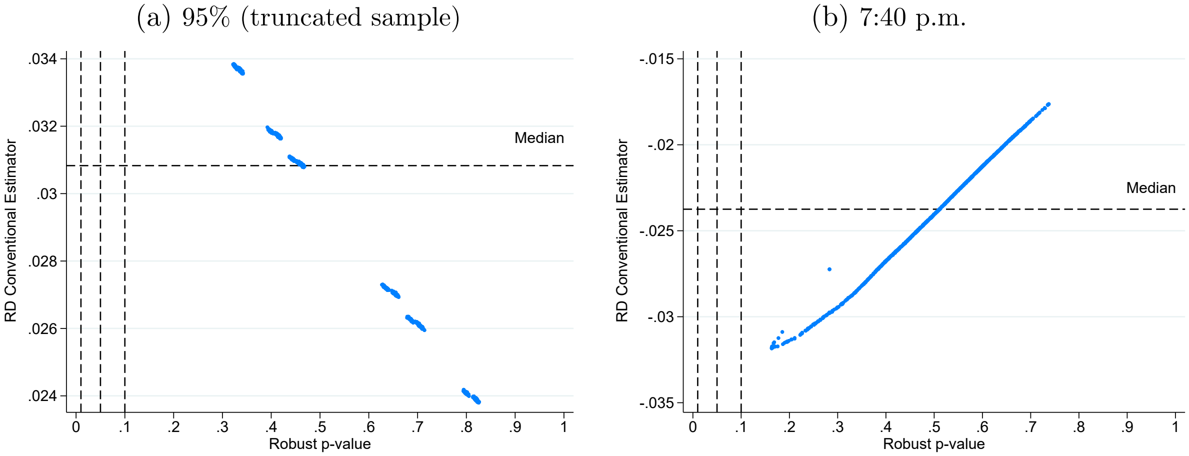
G RD estimate: Robustness

Sort order. As noted in the main text, only 8% of observations have unique time stamps. This is not surprising given the number of tally sheets and the length of the reporting window: there are 34,555 tally sheets, almost all of which were verified within a two-hour window, or 7,200 seconds (the time stamps include seconds, but not milliseconds). In the main text, we present results based on sorting the observations first by time stamp and then by a random number.

Of course, the sort order could affect our regression discontinuity (RD) results. To investigate whether our main RD result—failure to reject the null of continuity—is robust to different possible sort orders, we repeat the analysis 1,000 times, each time sorting (within time stamp) according to a different random draw. This exercise reveals that our failure to reject continuity is not the artifact of a specific sorting.

Figure G.1: No evidence of discontinuities, regardless of sort order

Each figure plots the magnitude of the RD estimate against the corresponding p-value, for each of 1,000 draws of the random variable used to sort observations within time stamps.


(a) 95% (truncated sample)

(b) 7:40 p.m.

Figure G.1 plots the magnitude of the RD estimates against the corresponding p-values for each of the 1,000 draws, for each of the three cutoffs studied in the paper. Table G.1 summarizes the results. The mean and median robust p-values are above 0.5, implying that the results presented in the main text are not anomalous: there is no evidence of a statistical discontinuity in MAS vote share at those cutoffs.

Table G.1: No evidence of discontinuities, regardless of sort order

Cutoff	Date & Time	Sample*	Robust p-value Mean	Robust p-value Median	RD Estimate Mean	RD Estimate Median	N Sortings
0.950	10/20/2019 20:03:59	Truncated	0.553	0.465	0.029	0.031	1,000
0.852	10/20/2019 19:40:57	Full	0.503	0.511	-0.024	-0.024	1,000

Polynomial degree and bandwidth. The results in the main text show that we cannot reject the null of continuity at the three cutoffs using a degree-one local polynomial with the MSE-optimal bandwidth. Table G.2 shows that, indeed, we cannot reject the null of continuity for other combinations of polynomial degree and bandwidth. Specifically, for each polynomial degree $p \in \{1, 2, 3\}$, we estimate the treatment effect using bandwidths selected with and without the regularization term (the regularization term shrinks the optimal bandwidth, Cattaneo, Idrobo and Titiunik, 2019, Section 4.4.2).

Table G.2: Robustness to polynomial degree and bandwidth choices

Cutoff	Date	Sample	Reg.	Deg.	Estimate	BW	p-val.	Robust C.I.	N Left	N Right
0.950	10/20/2019 20:03:59	Truncated	1	1	0.024	0.040	0.816	[-0.052, 0.065]	1,267	1,316
0.950	10/20/2019 20:03:59	Truncated	1	2	0.011	0.048	0.719	[-0.082, 0.057]	1,515	1,575
0.950	10/20/2019 20:03:59	Truncated	1	3	-0.010	0.057	0.204	[-0.129, 0.028]	1,823	1,662
0.950	10/20/2019 20:03:59	Truncated	0	1	0.028	0.092	0.365	[-0.042, 0.114]	2,926	1,662
0.950	10/20/2019 20:03:59	Truncated	0	2	0.009	0.078	0.685	[-0.103, 0.067]	2,485	1,662
0.950	10/20/2019 20:03:59	Truncated	0	3	-0.007	0.077	0.253	[-0.119, 0.031]	2,444	1,662
0.852	10/20/2019 19:40:57	Full	1	1	-0.028	0.044	0.374	[-0.077, 0.029]	1,455	1,457
0.852	10/20/2019 19:40:57	Full	1	2	-0.010	0.050	0.964	[-0.064, 0.061]	1,662	1,666
0.852	10/20/2019 19:40:57	Full	1	3	-0.010	0.069	0.891	[-0.072, 0.063]	2,260	2,286
0.852	10/20/2019 19:40:57	Full	0	1	-0.010	0.108	0.972	[-0.090, 0.087]	3,533	3,658
0.852	10/20/2019 19:40:57	Full	0	2	-0.024	0.081	0.583	[-0.065, 0.037]	2,644	2,701
0.852	10/20/2019 19:40:57	Full	0	3	-0.016	0.118	0.594	[-0.117, 0.067]	3,876	4,038

"Truncated" denotes the sample that excludes the voting booths without time stamps in the preliminary results system. "Reg." reports whether we choose the bandwidth with or without the regularization term; "Deg." reports the degree of the local polynomial.

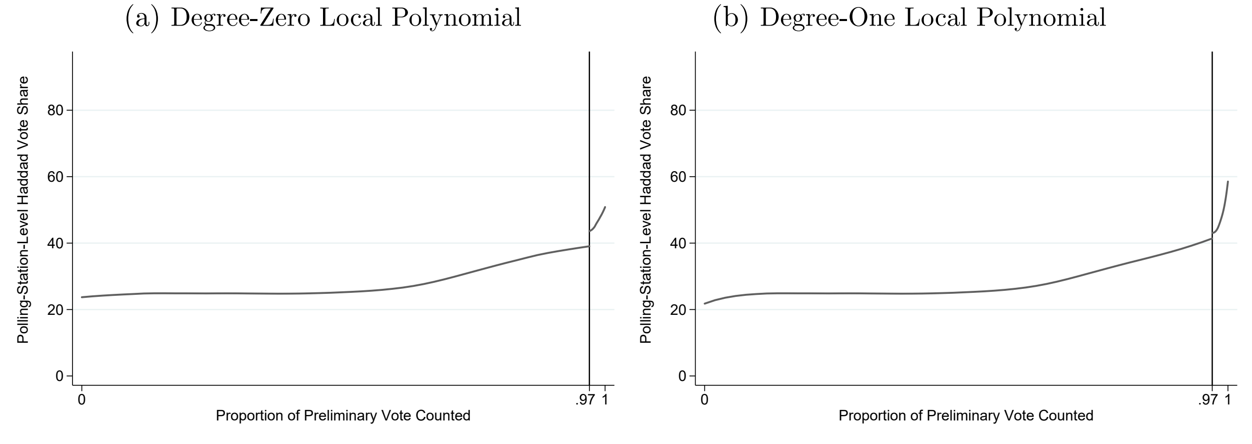
H Additional tables and figures

Figure H.1: Paper Ballot in Bolivia's Presidential Election

Source: Jorge Bernal

Figure H.2: Boundary Bias and Artificial Jumps in Haddad's Vote Share in Brazil

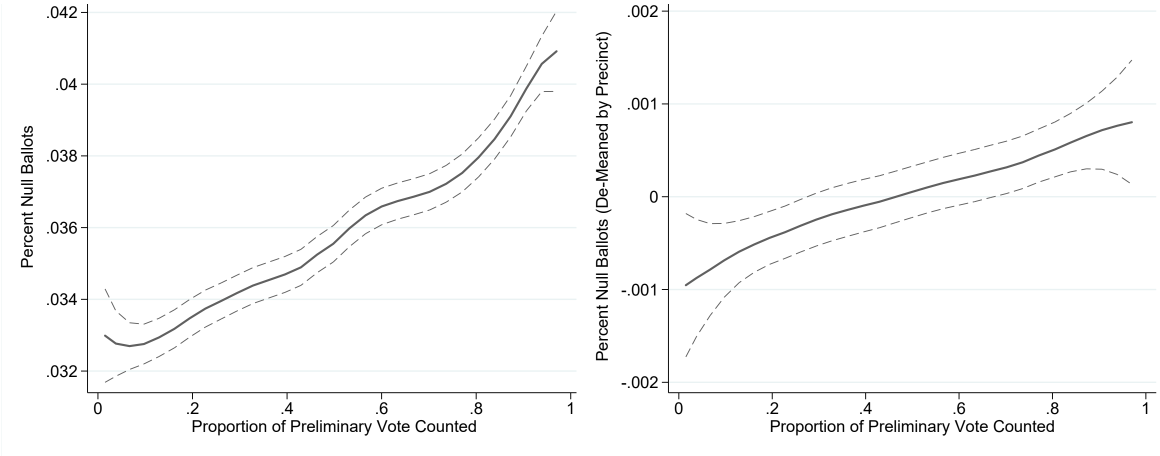
Figure (a) reveals that using a local constant fit creates the artificial appearance of a jump in Haddad's vote share at 97%. Figure (b) reveals that using a local linear fit corrects this.



Rule-of-thumb bandwidth from Fan and Gijbels (1996, p. 110–113), Epanechnikov kernel.

Figure H.3: Preliminary Results System Time is Correlated with % Null Ballots

Less-educated voters are more likely to cast null ballots. Consistent with the hypothesis that voting booths with less-educated voters were more likely to report later, the share of null ballots rises over the reporting window (a). And consistent with the hypothesis that within-precinct variation in socio-economic status drives within-precinct variation in reporting time, within-precinct variation in null ballot share is correlated with reporting time (b).

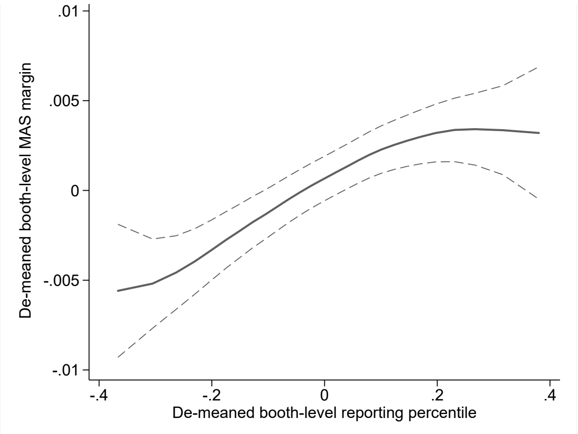

(a) % Null Rises Over Time

(b) Residual % Null Rises Over Time

Grey lines mark local linear fits using the rule-of-thumb bandwidth from Calonico, Cattaneo and Farrell (2018); the dashed lines mark 95% confidence intervals. Top and bottom 1% of de-meaned reporting times are excluded.

Figure H.4: Monotonicity of within-precinct variation

This plot shows the non-parametric relationship between de-meaned MAS vote margin ($m_{bp} - \overline{m}_p$) and de-meaned reporting time percentile (Time $\text{Percentile}_{bp} - \overline{TimePercentile_p}$).



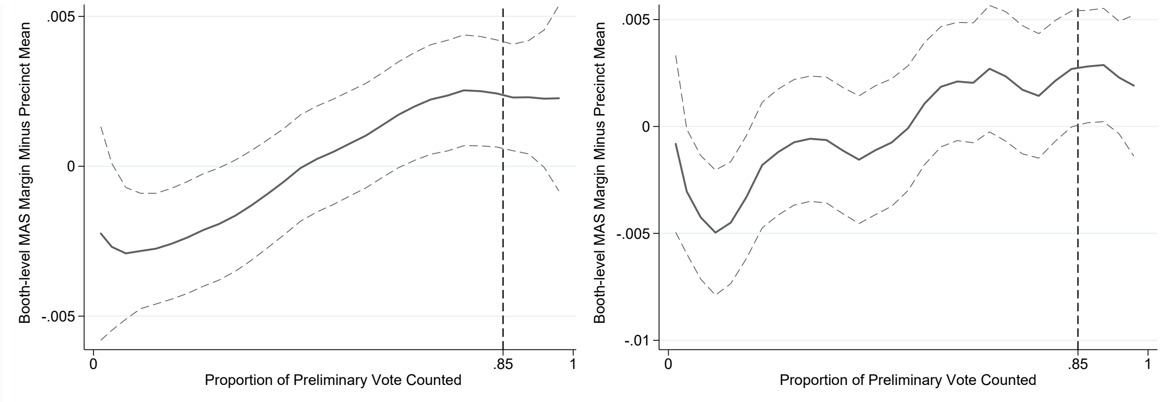
The grey line marks a local linear fit using the rule-of-thumb bandwidth from Calonico, Cattaneo and Farrell (2018); the dashed lines mark 95% confidence intervals.

Figure H.5: Within-Precinct Variation Trend, Smaller Bandwidth

Figure (a) repeats Figure E.1b; the takeaway is that, after accounting for precinct characteristics, the growth in MAS's margin does not accelerate after the public information blackout. Figure (b) shows that this result is not an artifact of bandwidth choice.

(a) Rule-of-thumb bandwidth (Calonico, Cattaneo and Farrell, 2018)

(b) Arbitrary small bandwidth (0.15)



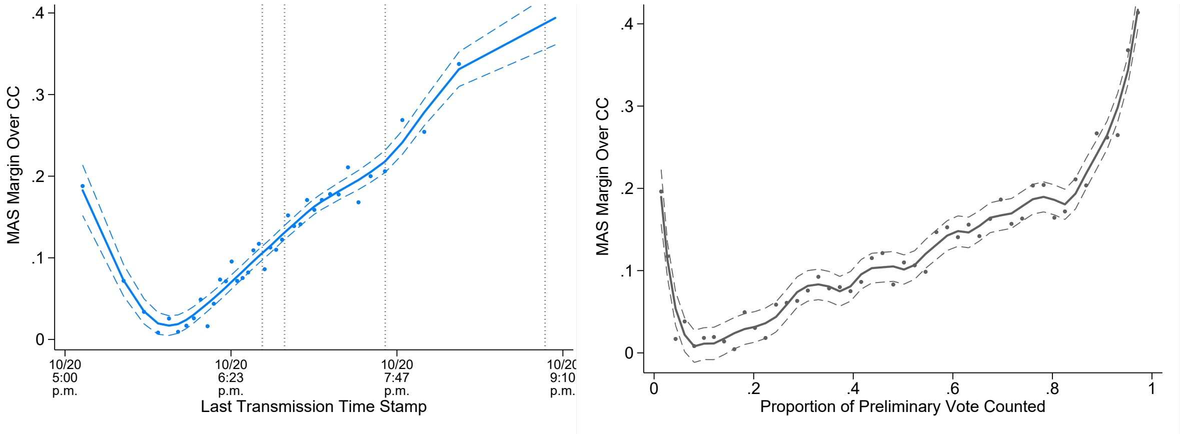
Grey lines mark local linear fits using the rule-of-thumb bandwidth from Calonico, Cattaneo and Farrell (2018); the dashed lines mark 95% confidence intervals.

Figure H.6: The Artificial Change in Slope

Figure (a) shows the time trend in Morales's vote share. Figure (b) shows this same trend as a function of *quantiles* of the vote counted. OAS (2019*a*) and Nooruddin (2020*c*) commented on "the steep slope" of this trend after approximately 80% of the vote had been counted. But this "steep slope" is simply an artifact of transforming the underlying time variable: as the optimal bins in (a) reveal, more time elapsed between the 87th and the 97th percentile than between the 47th and the 57th (dotted vertical lines). The transformation in (b) obscures this fact, artificially compressing time early and late in the evening.

(a) Margin vs. Clock Time

(b) Margin vs. Time Percentiles



Points mark the average MAS margin over Civic Community in optimal (data-driven) bins of the time variable (Cattaneo et al., 2019); lines mark local linear fits following Calonico, Cattaneo and Farrell (2018). Both figures trim the top and bottom 2% of observations; for a version without trimming, see Appendix Figure C.1a.