Information Versus Control:
The Electoral Consequences of Polling Place Creation[*]

Jeremy Bowles[†] & Horacio Larreguy[‡] & Anders Woller[§]

September 9, 2020

Abstract

We examine the incentives incumbents face when creating new polling places. First, doing so improves incumbents' ability to monitor brokers and voters by reducing the number of registered voters per polling station. Second, it reduces the distance traveled by citizens to vote, which undercuts incumbents' ability to control the electorate via turnout buying. We evaluate this trade-off in the context of Uganda, where the incumbent significantly influences electoral administration. Drawing on rich administrative data, we leverage discontinuities in the creation of polling places to causally identify the independent effects of (1) the number of voters per polling station and (2) distance to vote on electoral outcomes. We find that decreasing (1) improves incumbent electoral outcomes, while reducing (2) worsens them. The benefits for incumbents outweigh the costs, which rationalizes recent developments to expand polling infrastructure in Uganda and elsewhere.

Word count: 11,901

[†]PhD Candidate, Department of Government, Harvard University. jbowles@g.harvard.edu.

[‡]Associate Professor, Department of Government, Harvard University, and Visiting Researcher, Institute for Advanced Study in Toulouse and Toulouse School of Economics. hlarreguy@fas.harvard.edu and hlarreguy@tse-fr.eu.

[§]Postdoctoral Fellow, Department of Political Science, University of Copenhagen. an@ifs.ku.dk.

1   Introduction

Incumbent parties in developing countries engage in a multiplicity of strategies, some of which run afoul of democratic norms, to garner electoral support. A rich literature studies how these parties engage in clientelistic exchanges with voters (Finan and Schechter 2012; Nichter 2008; Stokes 2005) and even commit outright outright electoral fraud, including deliberate miscounting of votes or fraudulent registration (Birch 2012; Cantú 2014, 2019; Hidalgo and Nichter 2016; Simpser 2013). These strategies are furthermore conditioned by features of the electoral administration they face, including those affecting citizens' costs of voting (Bowles et al. 2019; Gans-Morse et al. 2014; Hariri et al. 2019).

While such features are typically considered fixed, incumbents often influence electoral administration (Chernykh and Svolik 2015). For instance, in many developing countries, incumbents may have a voice in deciding the location of electoral precincts (Wong 2019) and the resources allocated to them (Harris 2020). Some of these decisions directly affect voting behavior. This is particularly true of decisions over the creation and spatial allocation of polling places, which determines who has to travel far to vote and who does not (De Kadt 2019).

In response to changing demographics and international guidelines, large-scale reorganizations of voting infrastructure have taken place in sub-Saharan Africa over recent decades. Comparing the most recent elections in a country to the earliest election for which we have data, the increase in the number of polling places exhibits drastic variation across countries. This begs the question of whether the administration of voting has electoral implications and, consequently, whether such variation reflects incumbents deliberately trying to influence electoral administration to hold on to power.

We study the incentives that incumbent parties have to expand voting infrastructure by increasing the number of polling places. We posit that incumbents face a trade-off between better

information about local electoral support and continued *control* over the electorate. First, an increase in the number of polling places in an area reduces the number of registered voters per polling place. This is useful for assessing incumbent support and conditioning rewards and sanctions both to party brokers who work in particular localities (Bowles et al. 2019; Gingerich 2020; Larreguy 2013; Larreguy et al. 2016) and to voters in that area (Gottlieb and Larreguy 2020; Rueda 2017). Voters and brokers, anticipating these contingent effects, might be incentivized to exert more effort in support of the incumbent party (Rosenzweig 2019). Better information about local support, then, likely improves electoral outcomes for incumbents.

Second, an increase in the number of polling places may reduce the distance that citizens have to travel to register to vote and cast their ballot (De Kadt 2019). The resource advantage typically enjoyed by incumbents means they are likely to possess a comparative advantage in providing transport to mobilize likely supporters far from their polling places (Nichter 2008). This comparative advantage in turnout buying permits incumbents to control the *de facto* electorate by shaping the distribution of electoral access across space (Bowles et al. 2019; Larreguy et al. 2016). Since shorter distances to vote make it harder for the incumbent to effectively disenfranchise opposition voters, increasing the number of polling places in an area may worsen incumbents' electoral prospects by weakening their control over electoral access.

Depending on which effect dominates—the increase in information about local support, versus the reduction in control over electoral access—incumbents face incentives to either expand or contract the number of polling places. Whether they are *able* to do this depends on the extent of their influence over the administration of elections. This influence is likely to exist even in the presence of nominally independent electoral commissions, and is likely to be greatest in electoral autocracies (Chernykh and Svolik 2015; Simpser 2013).

To test this theoretical argument, we study the 2016 Ugandan presidential elections. As a

canonical electoral autocracy, the incumbent National Resistance Movement (NRM) party exercises substantial influence over the administration of elections (Tripp 2010). The Electoral Commission (EC), in turn, has repeatedly reorganized the electoral landscape with the official aims of decongesting polling places and shortening the distances voters must travel to cast ballots. We study what these reorganizations of the voting infrastructure reveal about the incentives of the incumbent party to sustain its electoral support.

To determine whether the information mechanism or the control mechanism dominates, we assess the effect of (1) the number of registered voters per polling station, and (2) the weighted average distance that citizens need to travel to their polling place, on electoral outcomes.[1] Our theoretical argument suggests that both of these explanatory variables are endogenous to the creation of new polling places. We therefore instrument for these explanatory variables using a regression discontinuity design that leverages the repeated application of similar administrative rules to reorganize polling places between the 2011 and 2016 elections. The first round of reorganization, which took place in 2013, created new polling places for voter registration in 2015 when a polling place from the 2011 election exceeded 900 registered voters.[2] In the second round, which took place once voter registration had concluded in 2015, polling places with more than 900 registered voters in the new 2015 voter register were split into multiple polling places.

The intuition behind our identification strategy is that, while each application of the administrative rule is likely to affect both (1) and (2), if the effect of (1) and (2) is different in 2013 than in 2015, then we have two independent sources of variation to instrument for their effects on electoral outcomes. Importantly, since the overlap of discontinuity samples is only partial,[3]

---

[1]A polling place is the venue for registration and voting, and can be split into multiple colocated polling stations.

[2]For Kampala and Wakiso—the two capital districts—the cutoff was 1,200 voters, but for simplicity we speak of a 900-voter cutoff throughout.

[3]If a polling place is split in the first reorganization, it is less likely that it will find itself close to the threshold again in the subsequent reorganization. A polling place may thus appear within one discontinuity sample but not the other.

we adopt seemingly unrelated regression-style estimating equations. Specifically, we stack the data of 2016 polling places within a bandwidth of the thresholds determining polling place splits in 2013 and 2015. To evaluate effects we use unusually rich administrative data obtained from various sources. These include the complete voter registers for the 2011 (n = 13.7 million) and 2016 (n = 15.3 million) elections as well as geocoded polling station-level results for both elections.

To gauge the magnitude of the trade-off faced by incumbent parties, we focus on the first stage of the instrumental variable estimation to assess the effect of a predicted additional polling place on the number of registered voters per polling station and the average distance to polling places in 2016. Together with the instrumental variable estimates, the first stage estimates then allow us to assess how an additional polling place affects turnout and vote share for the incumbent and opposition. Lastly, we focus on the reduced form effects to directly estimate the effect of an additional polling place on electoral outcomes.

Our first stage estimates are aided by the sequence of the two administrative rules, which affected the number of voters per polling station and their average distance to their polling places differently in 2013 than in 2015. Since the 2013 reorganization took place *prior* to voter registration, it increased the total number of registered voters by reducing the distance unregistered voters needed to travel to register. Our first stage estimation therefore accounts for how the effect of predicted polling place creation during the 2013 reorganization varies as a function of the distance of previously unregistered voters to their corresponding polling place before 2013. We find that predicted polling place creation in 2013 leads to a small average increase in distances faced by voters, which is driven by the registration of remote, previously unregistered voters. In contrast, for the reorganization that took place *after* registration in 2015, and thus had no effect on overall registration levels, predicted polling place creation has a strictly negative effect

on distance. Lastly, we find that predicted polling place creation both in 2013 and 2015 has a negative effect on the number of registered voters at affected polling stations in 2016.

Our instrumental variable estimates confirm the incumbent's trade-off. On the one hand, a one standard deviation reduction in the number of registered voters (a drop of 190 voters) leads to a 2 percentage point (pp) *increase* in incumbent, but not opposition, turnout. This effect is consistent with our hypothesis that splitting polling places increases the incumbent's ability to mobilize voters due to more precise information about voting behavior. On the other hand, a one standard deviation reduction in the distance of voters to their polling place (corresponding to 2 kilometers) leads to a 7pp *decrease* in vote share for the incumbent. This effect is driven mostly by increased opposition turnout, thus highlighting how the incumbent can exert control over the electorate by manipulating distances faced by voters.

Taken together, as the reduced form effects ultimately show, the effect of an additional polling place depends on its relative effects on the number of registered voters and the distance to vote, as well as how each of those variables affects electoral outcomes. The polling places that were created in the 2013 reorganization unequivocally benefitted the incumbent. Overall, this change improved the incumbent's vote share, especially so in more remote areas with high numbers of previously unregistered voters, with effects of up to 3pp. In contrast, additional polling places resulting from the 2015 reorganization overall weakly hurt the incumbent's vote share by less than 1pp. This outcome was the result of two effects: first, the reduction in the number of registered voters per polling station did not contribute much to incumbent turnout, and second, in contrast to the 2013 reorganization, an average reduction in the distance to vote increased opposition turnout. Back of the envelope calculations suggest that the creation of new polling places in 2013 and 2015 overall contributed to the incumbent's vote share in 2016, particularly among citizens who were far from polling places prior to 2013.

5

Finally, to probe the broader implications of these results, we compile an original dataset of the number of polling places in each presidential election in sub-Saharan Africa since 1990. We document substantial variation in the number of polling places by country and over time. These changes reflect differences in the motivations and abilities of different incumbent parties to expand voting infrastructure. We show that across sub-Saharan Africa, countries with *less* independent electoral commissions, which are characteristic of autocracies and countries with strong presidentialism, have *more* polling places for the same number of registered voters. This stylized fact is consistent with our results that more polling places favors the incumbent, suggesting that apparent increases in the representation of more marginalized citizens might be driven by motivations other than a sincere effort to eliminate barriers to electoral participation. It follows that our results carry practical importance by pointing to the potential partisan effects of a tool that is broadly believed to strengthen democracy.

2 Theoretical framework

An increase in the number of polling places in a territory is likely to simultaneously reduce the average number of registered voters per polling station and the distance that registered voters must travel to cast their vote. We argue that, in contexts where incumbents have influence over electoral administration and a comparative advantage in mobilizing voters through turnout buying, they face conflicting incentives to expand the number of polling places. While a reduction in the number of registered voters per polling station improves the information incumbents can obtain to monitor and hence incentivize electoral support, reductions in distance traveled to cast a vote might undermine their control over the composition of the electorate.

Information on levels of local electoral support

A growing literature studies the dynamics of voter mobilization in clientelistic democracies (Kitschelt and Wilkinson 2007; Stokes 2005). This literature studies the strategies that parties use to monitor, and subsequently incentivize, the behavior of both their political brokers and the groups of voters these brokers are tasked with mobilizing.

With respect to brokers, prior research has demonstrated how parties analyze electoral outcomes at the polling station-level, including turnout and vote shares, and construct noisy signals about brokers' performance on which they base rewards for brokers (Bowles et al. 2019; Gingerich 2020; Larreguy 2013; Larreguy et al. 2016).[4] Parties' ability to incentivize brokers' effort then depends on their capacity to monitor brokers' performance. This is easier to do when there are fewer voters per polling station. At small polling stations, electoral results therefore become more precise indicators of the performance of individual brokers (Bowles et al. 2019; Larreguy 2013; Larreguy et al. 2016). An improved capacity to monitor brokers' effort then allows parties to induce greater individual broker effort with respect to mobilizing voters and, consequently, attain better electoral outcomes.

With respect to voters, a related literature shows how parties use electoral results to condition the spatial targeting of rewards and sanctions. The creation of new, smaller polling places which map more closely to particular localities enables parties to extract better signals about their electoral support, which they can use to target distributive goods once in office (Gottlieb and Larreguy 2020). Additionally, because voters anticipate that their locality will be rewarded should their preferred candidate win, the availability of electoral results at low levels of aggregation should incentivize the turnout of particular groups aiming to signal their support (Rueda 2017). More precise information about electoral support in their area, therefore, may also in-

---

[4]Parties also monitor other outcomes of brokers' effort, such as turnout at rallies and political merchandising within their neighborhood (Szwarcberg 2014).

centivize voters to coordinate their vote (Jung and Long 2018).

The ability of a party to make use of this improved information depends on how well the party is equipped to mobilize turnout. Incumbent parties have access to more resources than challenger parties, which allows them to set up relatively sophisticated machines to monitor and incentivize their political brokers to ensure that they mobilize voters on their behalf (Bowles et al. 2019; Kitschelt and Kselman 2013). Therefore, more precise information about local electoral support is likely to incentivize citizens to vote for dominant incumbent parties in particular (Rosenzweig 2019).

Control of the electorate

Variation in citizens' costs of voting affects both their incentives to vote and their susceptibility to clientelistic modes of mobilization. A critical factor in citizens' decisions about whether to vote is the distance they need to travel to their polling place. Numerous studies from developed (Brady and McNulty 2011; Cantoni 2019; Haspel and Knotts 2005) and developing democracies (De Kadt 2019; Harris and van der Windt 2019) indicate that citizens' decisions to register to vote and turn out to vote are highly sensitive to the distance they need to travel.

Because citizens that need to travel further to their polling place face a higher cost of voting, these individuals are more likely to be influenced by turnout buying (Bowles et al. 2019; Gans-Morse et al. 2014; Larreguy et al. 2016). Given their significant resource advantages, incumbent parties are likely to possess comparative advantages in such turnout buying (Bowles et al. 2019; Kitschelt and Kselman 2013). In contrast, relatively under-resourced challenger parties do not have the means available to mobilize their likely supporters who face prohibitively long distances to their respective polling places.

In this way, high physical barriers to electoral access enable the incumbent to exert sig-

nificant control over the *de facto* electorate. Assuming reductions in distance affect incumbent

supporters *and* non-supporters alike,[5] the incumbent party has an interest in citizens facing high

costs of voting. By allowing the costs of voting to effectively disenfranchise citizens, the party

can selectively mobilize supporters who face high costs of voting, while leaving behind citizens

sympathetic to the opposition. The ability to selectively control the electorate weakens as costs

of voting are reduced.

Recent scholarship shows how parties respond to changes in the cost of voting by prioritiz-

ing vote buying over turnout buying (Gans-Morse et al. 2014; Hariri et al. 2019). Therefore,

reductions in the distance traveled to vote might still advantage incumbent parties, as they will

simply substitute from vote buying to turnout buying. However, in contrast to buying turnout,

monitoring whether voters vote as directed is difficult under the secret ballot, reducing the ef-

fectiveness of such vote buying vis-a-vis turnout buying (Nichter 2008). Consequently, from

the perspective of the incumbent, a residual advantage in vote buying is unlikely to compensate

for their reduced capacity to selectively mobilize voters.

Incentives to expand voting infrastructure

Incumbents often possess significant influence over electoral administration. Chernykh and

Svolik (2015) argue that electoral commissions with moderate pro-incumbent biases are likely

to exist in equilibrium in a broad set of developing democracies where incumbents possess su-

perior information about their electoral support than their challengers do. In contexts where

incumbents enjoy significant resource advantages and some degree of control over the admin-

istration of elections, incumbents face a trade-off when deciding whether to expand the number

---

[5]One could imagine that reductions in distance were intentionally implemented in supporting areas. However, directing reductions in the cost of voting specifically at supporters is challenging. This is especially so in areas with high returns to manipulation, such as highly competitive districts where precincts service equal numbers of supporters and non-supporters.

of polling places.

On the one hand, more polling places may lead to fewer registered voters per polling station, which helps the incumbent access more precise information about local electoral support. This improved ability to monitor brokers and voters may ultimately improve the incumbents' electoral performance. On the other hand, more polling places are likely to reduce incumbents' control over the electorate because the costs of voting decline, making it harder to selectively buy turnout. This is likely to have a negative effect on their electoral performance. An incumbent's incentive to increase or decrease the number of polling places is influenced by which of these two competing effects dominates.

Empirical implications

We summarize the empirical predictions of our theoretical framework here. Let β_1 represent the effect of (1) the number of registered voters in a polling station and let β_2 represent the effect of (2) the distance voters must travel, on electoral outcomes for the incumbent. Let π_1 and π_2 represent the effect of creating a new polling place on (1) and (2), respectively. Finally, let α represent the overall effect of creating a new polling place on electoral support for the incumbent.[6]

First, we formalize the expected effect of polling place creation on the number of registered voters per polling station, and the distance voters must travel to vote.

Hypothesis 1 *Polling place creation reduces the number of registered voters at each polling station and the distance faced by voters to cast their vote. Formally, $\pi_1 < 0$, $\pi_2 < 0$.*

Second, we formalize the conflicting incentives incumbents face.

---

[6]Mechanically, given our theoretical framework, $\alpha = \beta_1 \pi_1 + \beta_2 \pi_2$.

Hypothesis 2 *Increases in the number of registered voters in a polling station worsen electoral outcomes for the incumbent, while increases in distance to vote improve them. Formally, $\beta_1 < 0, \beta_2 > 0$.*

Third, we formalize the net effect of polling place creation and what it implies for the expansion of voting infrastructure.

Hypothesis 3 *If $\beta_1 \pi_1 + \beta_2 \pi_2 = \alpha > (<)$ 0, i.e., the net electoral effects of polling place creation are positive (negative), incumbents are (dis)incentivized to increase the number of polling places.*

3 Electoral politics in Uganda

We test our theoretical argument in the context of the 2016 presidential election in Uganda. The National Resistance Movement (NRM), led by president Yoweri Museveni, has held power continuously since 1986. Museveni was re-elected for his fifth term in office with 61% of the vote in 2016, defeating his long-time rival Kizza Besigye of the Forum for Democratic Change (FDC). Uganda is broadly classified as an electoral autocratic regime (Tripp 2010), with the incumbent party enjoying substantial resource advantages over the opposition (Platas and Raffler 2020). Below, we describe NRM's advantages in terms of the party's finances, its ability to mobilize voters, and its control of electoral administration.

Financial resource advantages

Political parties, especially the incumbent, spend a significant amount on elections (Grossman and Michelitch 2018). Running for office is expensive for candidates, with conservative estimates suggesting that candidates spent as much as $60,000 for a parliamentary seat (Collord

2016). Of the $716 million spent on the 2016 parliamentary and presidential elections, an overwhelming majority was spent by either NRM candidates or independents (ACFIM 2015).[7]

Much of these funds, primarily raised by candidates themselves (Collord 2016), support clientelistic initiatives. In a survey of more than 28,000 Ugandan respondents around the 2016 elections, 40% of respondents acknowledged receiving cash for their vote (Blattman et al. 2020). This share is virtually unchanged from the 2011 elections (Afrobarometer 2012). It is overwhelmingly candidates affiliated with the NRM that have the capacity to conduct this mobilization strategy (EOM 2016, p. 18), and of the respondents that reported receiving cash for their vote, 83% were approached by agents representing the incumbent party (Blattman et al. 2020).

Party organization and clientelistic strategies

Interviews with 438 brokers in Uganda, conducted after the 2016 elections, indicate that voter mobilization relies on the efforts of brokers, who mobilize supporters by arranging transport and by materially compensating supporters for their time (Blattman et al. 2020). An overwhelming share of village-level brokers are recruited locally. Brokers are usually well-connected individuals, who both know and are known by their fellow villagers.

Because brokers play such an important role, political parties exhibit a high level of organization in how they recruit and manage them. The candidates' campaign apparatuses are organized as sophisticated pyramid structures, with leaders at the constituency level, coordinators at the subcounty level, and managers at the parish level, who are ultimately responsible for recruiting and managing village-level brokers. Parish managers reward their brokers based on their performance, which they monitor in various ways. More than half of the brokers in the survey mention explicitly that candidates keep track of their individual performance. Brokers'

---

[7]Independent candidates are often those that ran unsuccessfully in NRM primary elections but remain loyal to the president.

performances are evaluated based on polling station-level results and campaign rally attendance, as well as the general visibility of the candidate in the community. A quarter of brokers reported that they were contracted to deliver particular vote margins for the candidate in their locality of operation.

In Uganda, turnout buying is important to the NRM, which draws a lot of its support from rural areas. In order to achieve favorable electoral results, just short of half of surveyed brokers report spending a substantial share of the funds they are allocated on the transportation of likely supporters. Brokers report how they direct a significant share of their expenditures on *boda boda* drivers, taxis, and fuel to transport voters to their polling station. As one broker stated, "I would put all this money only on transport because transport is so crucial" (Luwero, May 2017). Brokers also describe how they cross-check the voters' register to ensure voters do in fact comply by turning up at the polling station on election day.

Control of electoral administration

The Ugandan Electoral Commission (EC) maintains little independence from the NRM. Commenting on the 2016 elections in Uganda, the European Union's Election Observation Mission concluded that the Electoral Commission "lacked independence and transparency", specifically in the appointment of the EC's board members, who are selected by the president and approved by the NRM-dominated parliament. The EC's partiality for the incumbent party is an important reason why elections are considered as falling "short of international standards for the conduct of democratic elections at key stages" (EOM 2016).

Ample anecdotal evidence indicates how the EC tilts the playing field to favor the NRM. For example, there were widespread reports that election materials arrived late at a large number of polling stations in opposition strongholds such as Kampala and Wakiso (Newsweek 2016).

In some instances, voters had to wait for seven hours before election materials arrived, only to realize that presidential ballots were not included (The Guardian 2016).

The electorate is aware of the EC's partisanship. Citizens' trust in the EC is waning, with close to 55% reporting "little" or "no trust at all" in the EC in a 2018 survey, up from 40% in 2015 (Afrobarometer 2015, 2018). In the same survey of Ugandan citizens prior to the 2016 elections, around 20% believed that appointing a non-partisan EC was the most pressing issue to address.

# 4 Administrative reorganization of polling places

In recent years the electoral commission has repeatedly reorganized the administration of elections. We focus on two central components of the electoral administration in Uganda: the polling *place* and the polling *station*. A polling *place* is the venue for registration and voting, and it is often located at a school, market, or church. Polling places are split into multiple colocated polling *stations* where citizens actually cast their vote. Since Uganda did not maintain a permanent voter register at the time, all citizens seeking to vote in the 2016 election had to register in any of the polling places in their parish of residence or of birth. Once the National Voters' Register (NVR) was finalized, registered voters were assigned to the polling place closest to their village of residence (EC 2019).

Figure 1: Timeline of polling place reorganizations

Administrative rules determined two rounds of polling place reorganization between 2011

14

and 2016. These reorganizations officially aimed at enabling voters to conveniently cast their votes by reducing (1) the number of registered voters per polling station, and (2) the distance that voters needed to travel to their polling place (EOM 2016). Figure 1 provides a timeline. The first round of reorganization took place in August 2013 ("2013 reorganization"), *prior* to voter registration, and the second round took place from January to April 2015 ("2015 reorganization"), *after* voter registration. The 2013 reorganization split polling places exceeding 900 registered voters (or any such multiple) into several, geographically dispersed polling places on the basis of the 2011 voter register. The 2015 reorganization also split polling places or added extra polling stations to polling places exceeding 900 registered voters (or any multiple of 900) after the completion of voter registration.[8] Figure 2 illustrates the partial compliance with the 900-voter administrative rules determining polling place reorganizations, which motivates the fuzzy regression discontinuity design we explain below.

(a)   (b)
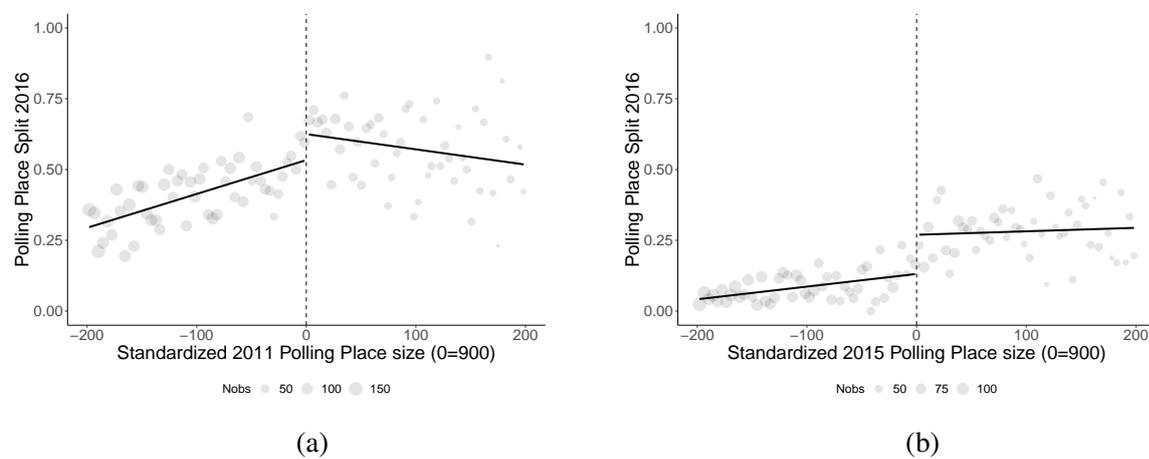
Figure 2: Compliance with the 900-voter administrative rule.
Note: The outcome in plots (a) and (b) is whether there was a polling place split (i.e. a new polling *place* was created).

---

[8]When splitting a polling place into multiple stations, voters were equally divided among the polling stations, generally according to their last names. For example, with two polling stations at the same polling place, one would serve voters with, say, last names beginning with A-K and the other last names beginning with L-Z.

Illustration of the effect on the number of registered voters and distance to vote

To illustrate how the polling place reorganizations worked, Figure 3 gives the example of how the 2013 and 2015 reorganizations affected the number of registered voters per polling station and distance to vote in the Kiboota parish in Western Uganda. In 2011, polling place *A* had 1,063 registered voters from five villages, who were on average 978 meters away. To comply with the 2013 reorganization, polling place *B* was added prior to voter registration. Once registration was over, *A* now had 964 registered voters from three villages, who were on average 808 meters away, while *B* had 512 registered voters from two villages, who were on average 848 meters away.

Figure 3: Reorganization of polling places in Kiboota parish

Note: Circles indicate polling place and triangles illustrate voters in villages. Color indicates village–polling station affiliation. Dashed line indicates parish border.

As a result of the 2015 reorganization, a third polling place *C* was added since *A* still exceeded 900 registered voters. Voters from the five villages in the parish were now distributed at three polling places for the 2016 election. The average polling place size was 490 (compared to

1,063 in 2011) and the average distance was 552 meters (compared to 978 in 2011). Our theoretical framework suggests that this change represents mixed news for the incumbent party. On the one hand, the improved mapping between villages (where brokers are also typically organized) and polling places improves what information the incumbent can extract from electoral results. On the other hand, with distance needed to travel to vote falling substantially, the incumbent party's ability to shape the electorate in the parish is weakened since it cannot—to the same extent—leverage its comparative advantage in voter mobilization.

Aggregate implications for the number of registered voters and distance to vote

The effect of the two reorganizations is also apparent at the national level. Even though the number of registered voters in Uganda increased between the 2011 and the 2016 elections—13.7 million voters registered in 2011 compared to 15.3 million in 2016—the 2013 and 2015 reorganization reduced the number of registered voters per polling station, as illustrated in Figure 4.
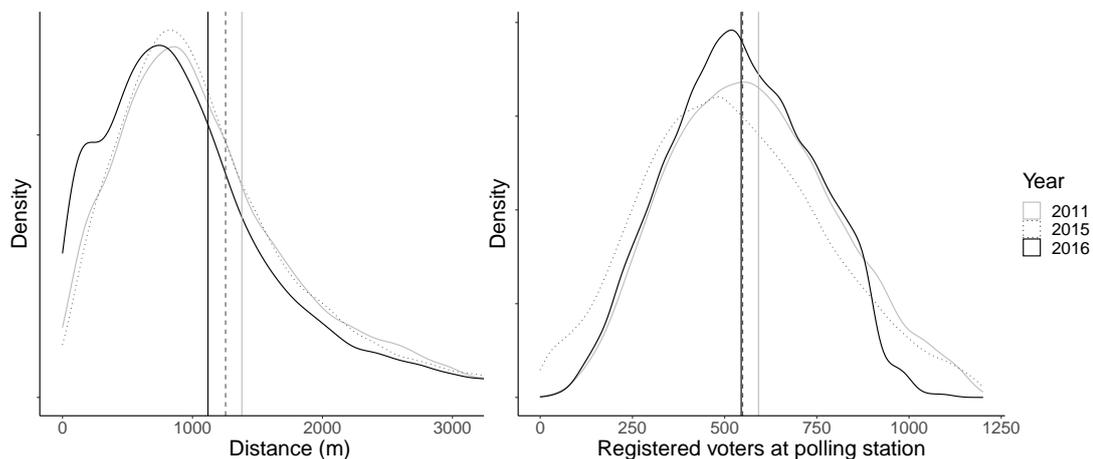

Figure 4: Distribution of distance to polling place and registered voters at polling station
Note: Horizontal lines indicate the mean. Distance is weighted by number of registered voters.

In turn, Figure 4 masks two countervailing effects that the reorganizations had on the distance that registered voters had to travel to vote. The 2013 reorganization was implemented prior

to voter registration for the 2016 election. Because it created new, geographically dispersed polling places, the 2013 reorganization had the effect of increasing the number of registered voters by making registration easier. The average distance of registered voters to their polling place sometimes actually *increased* after the 2013 reorganization when individuals who lived in remote parishes and previously were not registered due to long distances now were registered.[9] The 2015 round of reorganization, in contrast, was implemented after voter registration was concluded, thus it had no effect on overall voter registration numbers. It did, however, create new, geographically dispersed polling places to serve registered voters, and thus the average distance of registered voters to their polling place went *down* after the 2015 reorganization.

This variation in the effects of the two rounds of reorganization on distance to polling places, as well as on the number of registered voters, allows us to estimate the effects of each variable on electoral outcomes. Below we describe how we use this variation for identification.

5 Data

We make use of three sources of administrative data to examine the electoral consequences of polling place creation. First, we use the complete voter registers for the 2011 elections (n=13.7 million) and 2016 elections (n=15.3 million) scraped from the Electoral Commission website. This data provides every voter's name, age, village, and polling station. Second, we use geolocated polling station–level presidential election results from 2011 and 2016. Third, we use administrative data sources, including geolocated data from the 2014 National Population and Housing Census and mappings of administrative units over time from the Electoral Commission of Uganda.

---

[9]Note that we are referring to a measure of average distance of *registered voters* to their polling places, not average distance of *all citizens* to their polling place. The former went up after the 2013 reorganization; the latter almost certainly went down.

Matching polling stations over time

Implementing our identification strategy requires that we match each polling *station* in 2016 to each polling *place* as it existed before the 2013 reorganization and as it existed in between the 2013 and 2015 reorganization. Doing this allows us to identify polling stations in 2016 that were created as a result of the 2013 or 2015 reorganizations. This data-intensive exercise, described fully in Appendix A.1, was accomplished through an iterative process of (1) standardizing the geographical identifiers of polling stations across 2013, 2015, and 2016; (2) matching based on any time-invariant polling station names; (3) matching based on high shares of the same registered voters in the 2011 and 2016 voter registers; and (4) matching based on high shares of voters from the same villages between the 2015 village-polling station data and the 2016 voter register.[10]

As a result, for each polling station in 2016, we identify its (potentially invariant) polling place before the first reorganization in 2013, before the second reorganization in 2015, as well as the corresponding number of registered voters and average distance to the polling place (in both 2013 and 2015), and prior electoral outcomes from 2011.

Our data is at the polling station–level in 2016, where we observe 28,010 polling stations nested within 24,355 unique polling places (up from 23,054 and 20,905, respectively, in 2013 and 2011). We provide descriptive statistics for the variables outlined below in Table A2.

Dependent variables

At the polling station–level in 2016, we construct outcomes measuring electoral support in the presidential election for the incumbent NRM party as well as aggregate support for all opposition

---

[10]We do not have a full voter register from after the 2015 registration period but before the 2015 reorganization. Such a dataset would comprise an identical set of voters, but they would be assigned differently across polling places. Instead, we have data from after the 2015 registration period which lists the number of voters per village at each polling station.

parties. For both, we consider (1) party turnout, defined as total votes received by the party divided by the total number of registered voters at that polling station; (2) party vote share, where the denominator is instead the total number of votes cast at the that polling station. We also calculate total turnout at each polling station (measured as all votes cast divided by total number of registered voters).

Independent variables

Our primary independent variables are the number of registered voters assigned to each polling station within a polling place and the weighted average distance that registered voters would face to cast their ballots.

The number of registered voters assigned to each polling station in the 2016 election is reported in the administrative data. The mean number of registered voters at a polling station was 592 voters in the 2011 election, 549 voters after the 2013 reorganization, and 545 voters in the 2016 election. To compute the weighted average distance of registered voters from their assigned polling place, we first geocode villages listed in the voter register as citizens' home localities ($n = 58,000$) by matching them to the 2014 village census containing the location of all villages in Uganda. For a given polling place, we then calculate its distance to each of the villages where its registered voters reside. We take the average of these distances, weighted by the numbers of registered voters per village, as our measure of average distance to vote at the polling place–level (for detailed information see Appendix A.2). Consistent with the expansion of the number of polling places over time, the weighted mean distance to a polling place was 1.46 km in the 2011 election, 1.33 km after the 2013 reorganization, and 1.28 km in the 2016 election.[11]

---

[11]The overall decline in distance following the 2013 reorganization is consistent with a selective positive effect of the same reorganization on distance. While many polling places experienced reductions in distance between the 2011 election and the period following voter registration, they are typically excluded from the subset comprising

6 Empirical strategy

In this section we outline an empirical strategy that permits the estimation of both the net effect of polling place creation on electoral outcomes, as well as the individual effects of the number of registered voters per polling station and average distance to the polling place.

Intuition

To estimate the effects of the number of registered voters and the average distance that voters have to travel to their polling place, consider the following 'structural' equation:

$$y_i = \beta_1 RegVoters_i + \beta_2 Distance_i + \epsilon_i, \tag{1}$$

where y_i measures an electoral outcome at polling station i. The key implication of our theoretical framework, captured by Hypothesis 2, is that the incumbent may face a trade-off in the creation of new polling places because $\beta_1 < 0$ and $\beta_2 > 0$ when y_i measures electoral support for the incumbent. Smaller polling places may be easier to monitor, but reducing the distance voters have to travel to vote could undercut the incumbent's comparative advantage in mobilization.

Estimating Equation (1) requires at least two instrumental variables for the two independent variables to causally estimate the β parameters. The first inferential challenge is that the number of registered voters at the polling station and the distance that they have to travel to their polling place are likely endogenous: confounders may determine both the allocation of new polling places and electoral outcomes.

To deal with this challenge, we use a fuzzy regression discontinuity design leveraging the discontinuous creation of new polling places between 2011 and 2016. Polling places existing

---

our discontinuity analysis sample of polling places affected by the reorganizations. Within this sample, as we show below, the effect of the 2013 reorganization on distance is weakly positive.

prior to the 2013 or 2015 reorganizations with *more than* 900 registered voters were more likely to be split into multiple polling places for the 2016 election than were polling places with *just below* 900 registered voters. Moreover, since the 2013 reorganization took place before voter registration, it increased the total number of registered voters by reducing the distances faced by previously unregistered voters. In turn, the timing of the reorganizations likely influences the effect of polling place creation on the number of registered voters at the polling station and average distance. While the 2015 reorganization reduced distance generally, the 2013 reorganization's effect on average distance is conditioned by the distance of previously unregistered voters, who were encouraged to register to vote. This implies two first stage specifications of the form:

$$y_{ip}^{13} = \alpha_1 T_p^{13} + \alpha_2(T_p^{13} \times DPUV_p^{13}) + \alpha_3(\text{Forcing}_p^{13}) + \alpha_4(T_p^{13} \times Forcing_p^{13}) + \alpha_5 DPUV_p^{13} +$$

$$+ \alpha_6(Forcing_p^{13} \times DPUV_p^{13}) + \alpha_7(T_p^{13} \times Forcing_p^{13} \times DPUV_p^{13}) + \omega X_p + \epsilon_{ip}^{13} \quad (2)$$

$$y_{ip}^{15} = \alpha_1 T_p^{15} + \alpha_2(Forcing_p^{15}) + \alpha_3(T_p^{15} \times Forcing_p^{15}) + \omega X_p + \epsilon_{ip}^{15} \quad (3)$$

where y_{ip}^Y indicates electoral outcomes at polling station i in 2016 which had been part of polling place p in $Y \in \{2013, 2015\}$. T_p^Y is an indicator for whether p contained more than the threshold number of registered voters (and therefore was likely to be split) in reorganization Y. $Forcing_p^Y$ is the forcing variable measuring the deviation between the number of registered voters in p in Y and the closest threshold (polling places are split at each multiple of 900 voters). $DPUV_p^{13}$ is the distance-weighted measure of previously unregistered voters to their polling place, which we standardize at mean zero and standard deviation one so that the coefficient on predicted polling station split captures the mean effect.[12] Last, X_p is a vector of pre-treatment

---

[12]We provide full robustness tests where we exclude this interaction term. However, as our first stage results indicate, doing this risks introducing a violation of the monotonicity assumption required for the instrumental variable estimates to be valid.

covariates defined at the polling place level prior to the 2013 reorganization, that includes total turnout, incumbent turnout, opposition turnout, and distance to vote. Standard errors are clustered at the level of the pre–2013 reorganization polling place. We restrict the sample to polling stations within a narrow bandwidth of the 900-voter threshold. Specifically, we define our baseline bandwidth as +/- 200 registered voters and demonstrate robustness to different bandwidth choices.

These two first stage specifications provide us with three candidate instruments T_p^{13}, T_p^{15}, and $T_p^{13} \times DPUV_p^{13}$. While this generates plausibly exogenous variation in the creation of new polling places, there is a potential second inferential challenge. Specifically, Equation (1) might be underidentified if the variation captured by the the three candidate instruments is not sufficiently independent to instrument for the two endogenous variables.

We address this potential challenge with the observation that, while the 2013 and 2015 reorganizations likely affected both independent variables in Equation (1) simultaneously, their effects likely differ both in magnitude and possibly sign due to differences in their implementation. This variation permits us to instrument for both explanatory variables using T_p^{13}, $T_p^{13} \times DPUV_p^{13}$, and T_p^{15} and hence identify β_1 and β_2 in Equation (1). So long as T_p^{13} and $T_p^{13} \times DPUV_p^{13}$ have sufficiently different relative effects on the first stage outcomes compared to T_p^{15}, the variation in these first stage effects may be leveraged to identify the independent effect of the number of registered voters in a polling station and the effect of distance to a polling place on electoral outcomes.

Estimating equations

The potential use of T_p^{13}, $T_p^{13} \times DPUV_p^{13}$, and T_p^{15} as instruments for polling station size and distance faces one more empirical challenge. Since our unit of observation is polling station i

in 2016, we effectively have two discontinuity samples: the subset of i that were part of 2011 polling places that were close to the threshold number of registered voters that determined a polling place split during the 2013 reorganization, and the (possibly overlapping) subset of i that were part of polling places after voter registration in 2015 that were close to the threshold number of registered voters that determined a polling place split during the 2015 reorganization. As Figure A2 indicates, these two discontinuity samples only partially overlap. We address this final challenge through implementing estimating equations in the spirit of a seemingly unrelated regression (SUR). The SUR setup stacks these two discontinuity samples together to account for the partial extent of their overlap.

Denoting the relevant discontinuity sample with superscripts, our outcome takes the form of the appended vector $(y_{ip}^{13}, y_{ip}^{15})$, which will have length $(n_{\text{BW}}^{13} + n_{\text{BW}}^{15})$ depending on the bandwidth, BW, used to define the two discontinuity samples and hence the number, n, of polling stations in 2016 that were part of polling place p in 2013 or 2015 within this bandwidth at the point of reorganization.[13] Similarly, the two endogenous variables are defined by the vectors $(RegVoters_{ip}^{13}, RegVoters_{ip}^{15})$ and $(Distance_{ip}^{13}, Distance_{ip}^{15})$. Finally, the three excluded instruments take the form of three vectors: $(T_p^{13}, 0)$, $(T_p^{13} \times DPUV_p^{13}, 0)$, and $(0, T_p^{15})$, with parallel forms for their relevant forcing variables and their interactions with the instruments.

The full estimating equation, in our baseline specification, therefore takes the following form:

$$\begin{pmatrix} y_{ip}^{13} \\ y_{ip}^{15} \end{pmatrix} = \alpha_1 \begin{pmatrix} T_p^{13} \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} T_p^{13} \times DPUV_p^{13} \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ T_p^{15} \end{pmatrix} + \alpha_4 \begin{pmatrix} Forcing_p^{13} \\ 0 \end{pmatrix} + \alpha_5 \begin{pmatrix} 0 \\ Forcing_p^{15} \end{pmatrix}$$
$$+ \alpha_6 \begin{pmatrix} Forcing_p^{13} \times T_p^{13} \\ 0 \end{pmatrix} + \alpha_7 \begin{pmatrix} 0 \\ Forcing_p^{15} \times T_p^{15} \end{pmatrix} + \alpha_8 \begin{pmatrix} DPUV_p^{13} \\ 0 \end{pmatrix} \quad (4)$$
$$+ \alpha_9 \begin{pmatrix} Forcing_p^{13} \times DPUV_p^{13} \\ 0 \end{pmatrix} + \alpha_{10} \begin{pmatrix} T_p^{13} \times Forcing_p^{13} \times DPUV_p^{13} \\ 0 \end{pmatrix} + \omega X_p + \epsilon_{ip}$$

---

[13]Throughout the estimation and analysis, we apply the same bandwidths to each discontinuity sample.

$$\begin{pmatrix} y_{ip}^{13} \\ y_{ipd}^{15} \end{pmatrix} = \beta_1 \begin{pmatrix} \widehat{RegVoters}_{ip}^{13} \\ \widehat{RegVoters}_{ip}^{15} \end{pmatrix} + \beta_2 \begin{pmatrix} \widehat{Distance}_{ip}^{13} \\ \widehat{Distance}_{ip}^{15} \end{pmatrix} + \beta_3 \begin{pmatrix} Forcing_p^{13} \\ 0 \end{pmatrix} + \beta_4 \begin{pmatrix} 0 \\ Forcing_p^{15} \end{pmatrix}$$

$$+ \beta_5 \begin{pmatrix} Forcing_p^{13} \times T_p^{13} \\ 0 \end{pmatrix} + \beta_6 \begin{pmatrix} 0 \\ Forcing_p^{15} \times T_p^{15} \end{pmatrix} + \beta_7 \begin{pmatrix} DPUV_p^{13} \\ 0 \end{pmatrix} \quad (5)$$

$$+ \beta_8 \begin{pmatrix} Forcing_p^{13} \times DPUV_p^{13} \\ 0 \end{pmatrix} + \beta_9 \begin{pmatrix} T_p^{13} \times Forcing_p^{13} \times DPUV_p^{13} \\ 0 \end{pmatrix} + \omega X_p + \epsilon_{ip}$$

where Equation (4) provides first stage and reduced form estimates, and Equation (5) is the instrumental variables specification. While we use a bandwidth of +/- 200 registered voters to define the discontinuity samples for the two reorganizations, we demonstrate robustness to different bandwidth choices.[14]

Identification assumptions

Several identifying assumptions must be satisfied to uncover the structural parameters of interest from Equation (1) by instrumenting $RegVoters_i$ and $Distance_i$ with T_p^{13}, $T_p^{13} \times DPUV_p^{13}$, and T_p^{15}. First, for relevance, we need that each of the instruments affect at least one of the two endogenous variables. Conditional on this, we need sufficient variation in the relative effects of the instruments on the two endogenous variables to be able to separately identify the structural parameters β_1 and β_2. The partial F-statistics from the first stage regressions and the F-statistic drawn from the two stage least squares (2SLS) estimation of Equation (5) discussed later generally support this assumption.

Second, exogeneity of the instruments requires that polling places that are just above the threshold in 2013 or 2015 are on average similar to those that are just below, as well as along the various values of $DPUV^{13}$. In Table A3 we regress a set of pre-treatment variables defined

---

[14]This bandwidth gives us a total of 12,895 observations, evenly split across the two discontinuity samples.

in 2013 onto Equation (4). Interpreting the balance tests, we find no systematic evidence of imbalance on pre-treatment measures of electoral outcomes, distance to vote, or geographical location for T^{13}, $T^{13} \times DPUV^{13}$, or T^{15}. Similarly, Figure A3 indicates that there are not differential densities of polling places around the discontinuities.

Third, the clear threat to monotonicity relates to the 2013 reorganization, since its timing before voter registration in 2015 creates the possibility that the distance to vote in some polling places *increases* due to inducing the registration of a large number of previously remote individuals. This concern is addressed by the inclusion of the interaction with $DPUV^{13}$, which enables us to isolate how the effects of polling place splits vary with the distance of previously unregistered voters. Fourth, the key potential threat to the exclusion restriction relates to the non-randomness of the location of the new polling places, even if their creation is as-if random. Table A4 shows that polling stations in 2016 included in any of the discontinuity samples, but just above or below the 900-voter threshold appear broadly identical in terms of local characteristics, which supports the plausibility of the exclusion restriction.

7 Results

First stage

First, we estimate how predicted polling place splits as a result of the two reorganizations affect the two endogenous variables of interest— the number of registered voters assigned to a polling station and their average distance to vote in 2016. Table 1, estimated using Equation (1), presents the first stage effects of a predicted polling place split on standardized measures of the number of registered voters at the polling station and distance to vote in 2016.

Interpreting the distance results first, we find that a predicted split in 2015 is associated with

Table 1: First stage

	Registered voters	Distance
	(1)	(2)
T^{13} (α_1)	-0.464***	0.070*
	(0.043)	(0.040)
$T^{13} \times DPUV^{13}$ (α_2)	0.024	0.163**
	(0.051)	(0.065)
T^{15} (α_2)	-0.485***	-0.146***
	(0.043)	(0.050)
Partial F-statistic	54.79	9.88
Observations	12895	12895

Unit of observation is the polling station in 2016. Outcome variables: Column (1) reports standardized number of registered voters per polling station; column (2) reports standardized distance of voters to polling station. Controls include turnout, incumbent turnout, opposition turnout, and distance to the polling place, all measured using data from 2011 election. Partial F-statistic computed following Sanderson and Windmeijer (2016).

All specifications are estimated using OLS (Equation 4) within a pooled bandwidth of +/-200 registered voters. Standard errors clustered at the 2013/2015 polling place-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

a 0.15 standard deviation (sd) decrease in the average distance voters had to travel to cast their votes ($p < 0.01$). In contrast, a predicted split in 2013 is associated with a weakly positive 0.07 sd *increase* in the distance to vote ($p < 0.1$). As indicated by Figure 5, which shows how the effect of a predicted split varies with the distance of previously unregistered voters, the positive effect on distance is explained by the previously unregistered voters living far from the new polling place. When previously unregistered voters were already proximate to the new polling place, the effects of a predicted split mimic the effects in 2015 with a reduction in distance.

Figure 5: First stage effects by the distance of previously unregistered voters

90% and 95% confidence intervals shaded in gray.
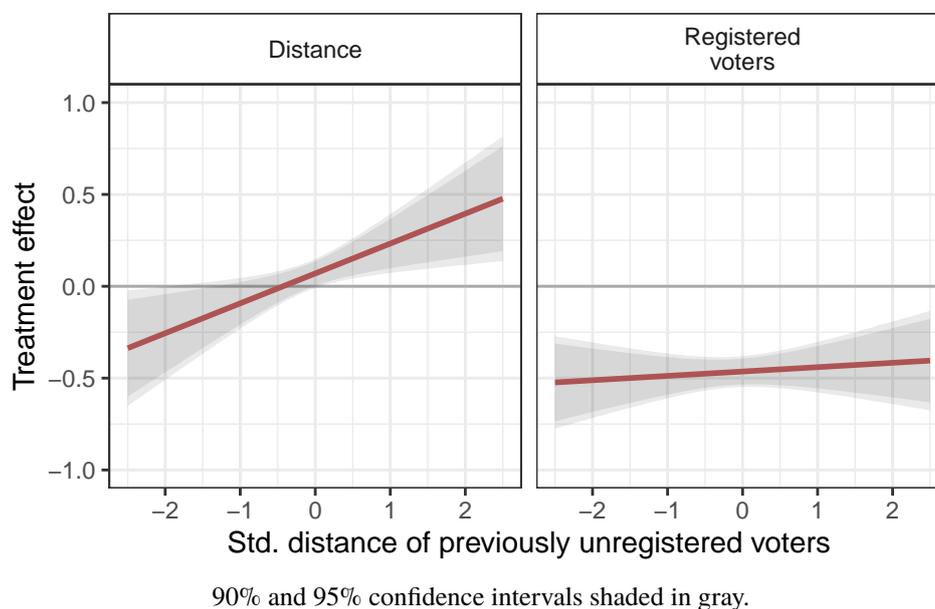
Turning to the effects of a predicted split on the number of registered voters, we find similar effects across the two rounds of reorganizations. We find that a predicted split in 2013 is associated with a substantively large 0.46 sd decrease in the number of registered voters at the corresponding polling station in 2016 ($p < 0.01$). A predicted split in 2015 generates a comparable 0.49 sd decrease in the number of registered voters at the corresponding polling station in 2016 ($p < 0.01$).

Overall, we find the partial F-statistics of the instrumental variables for both endogenous

regressors to be relatively high, with $F = 55$ for registered voters and $F = 10$ for distance.[15] As described above, this sufficiently independent variation in the first stages permits identification of the β_1 and β_2 parameters.

Instrumental variables

Next, we estimate Equation (5) to examine the independent causal effects of the number of registered voters per polling station and distance to vote on electoral outcomes. Table 2 presents the results for the five dependent variables of interest: (1) incumbent turnout; (2) incumbent vote share; (3) opposition turnout; (4) opposition vote share; and (5) total turnout.

Table 2: Instrumental variables estimates

	Incumbent		Opposition		
	Turnout (1)	Vote share (2)	Turnout (3)	Vote share (4)	Turnout (5)
Registered voters $(\widehat{\beta_1})$	-0.019**	-0.007	-0.002	0.010	-0.023***
	(0.009)	(0.011)	(0.007)	(0.011)	(0.007)
Distance $(\widehat{\beta_2})$	0.033	0.071**	-0.039**	-0.075**	-0.004
	(0.024)	(0.031)	(0.019)	(0.032)	(0.018)
Weak Instrument F-statistic	6.04	6.04	6.04	6.04	6.04
Control Mean	0.40	0.58	0.25	0.37	0.69
Observations	12895	12895	12895	12895	12895

Unit of observation is the polling station in 2016. Controls include turnout, incumbent turnout, opposition turnout, and distance, all measured in 2011.

All specifications are estimated using 2SLS (Equation 5) within a pooled bandwidth of +/-200 registered voters. Standard errors clustered at the 2013/2015 polling place-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

First, the estimated coefficient for registered voters (β_1) in Column 1 implies that a 1 sd increase in the number of registered voters assigned to a polling station (an addition of 190 voters) generates a 1.9 percentage point (pp), or 5%, decrease in turnout for the incumbent NRM party ($p < 0.05$). In contrast, the number of registered voters has no effect on overall

---

[15]We compute these partial F-statistics following Sanderson and Windmeijer (2016).

turnout for the opposition (Column 3). This suggests that the incumbent particularly benefits from the improved information associated with smaller polling stations. We do not interpret the effects of registered voters on party vote shares because of Column 5, which shows that a 1 sd increase in the number of registered voters in a polling station reduces total turnout by 2.3 pp ($p < 0.01$). As a result, the estimated effects on party vote shares are likely to be post-treatment biased.

Second, turning to the effects of distance, we find that a 1 sd increase in distance (an increase of 2 kilometers) increases the vote share received by the incumbent by 7.1 pp (12%), and reduces the vote share received by opposition parties by 7.5 pp (19%), both significant at the 5% level.[16] Judging from the weak positive effect of distance on incumbent turnout (Column 1), and the stronger negative effect on opposition turnout (Column 3), the effect of distance on vote shares is likely caused by opposition voters being effectively disenfranchised by the longer distances. The estimates in Column 5 relating to total turnout are consistent with this interpretation, with a 1 sd increase in distance having a relatively precise null effect on turnout.

The overall F-statistic for the IV analysis, which we compute following Kleibergen and Paap (2006), is around 6. While the conventional heuristic for a strong instrument is that such statistics should be at least 10, this applies only to the case of a single endogenous regressor (Staiger and Stock 1997). With multiple endogenous regressors, as we have, we therefore compare the F-statistic to the critical values provided by Stock and Yogo (2005).[17] This exercise indicates that we can reject levels of bias relative to the bias of OLS estimates of above 15%, which represents relatively strong identification.[18]

---

[16]Since the number of registered voters affects total turnout, the estimates in Columns 2 and 4 using vote shares are post-treatment biased and hence difficult to interpret. However, because distance has a null effect on total turnout (Column 5), we can consider the effects of distance on vote share for the incumbent and opposition.

[17]The Stock & Yogo critical values rely on the assumption of i.i.d. errors, which is violated in our setting since we cluster the standard errors. However, since alternative critical values are only computable in the case of one endogenous regressor, it is common empirical practice to nonetheless compare computed F-statistics to the Stock & Yogo values (Andrews et al. 2019).

[18]Two additional pieces of evidence suggest that weak identification is not significantly affecting our estimates.

The instrumental variables estimates confirm that the incumbent faces a trade-off. Creating new polling places improves their electoral outcomes through increased monitoring—the incumbent's turnout is greater at polling stations with fewer registered voters—while simultaneously worsening their electoral outcomes by reducing their control over the electorate. Because the incumbent has a comparative advantage in turnout buying, their electoral support is greater when voters face longer distances to vote.

Which effect dominates?

Finally, we consider the net effect of polling place creation on electoral outcomes. In Table 3, we report estimates of reduced form effects by regressing the five outcome variables in Table 2 onto the regressors in Equation (4). The results show that the 2013 reorganization contributed unequivocally to the incumbent's electoral outcomes in 2016. This is because the reorganization decreased the number of voters per polling station, which boosted the incumbent's turnout, while the average increase in distance to vote reduced the opposition's turnout. Overall, as Figure 6 indicates, splits in 2013 contributed to increases in the incumbent's vote share, especially so in places where previously unregistered voters were most distant, with effects of up to 3pp, or 5%.

In turn, the 2015 reorganization had an overall null effect on incumbent turnout, a weakly positive effect on opposition turnout, and a positive effect on total turnout. As a result, this leads to statistically insignificant drop in incumbent vote share of less than 1pp. This results from a combination of the reduction in the number of registered voters per polling station not contributing as much to incumbent turnout and, in contrast to the 2013 reorganization, an average reduction in distance to vote that increased opposition turnout.

First, we perform robustness tests where, at the risk of introducing a violation of the monotonicity assumption, we exclude the interaction $T^{13} \times DPUV^{13}$, which strengthens the overall F-statistic and has little effect on the IV estimates (see Table A8). Second, in Table A10 we estimate the "naive" OLS regression and find these estimates to be strikingly similar to the IV estimates.

Table 3: Reduced form estimates

	Incumbent		Opposition		
	Turnout	Vote share	Turnout	Vote share	Turnout
	(1)	(2)	(3)	(4)	(5)
T^{13} (α_1)	0.012**	0.009	-0.001	-0.011*	0.013***
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)
$T^{13} \times DPUV^{13}$ (α_2)	0.002	0.010	-0.008*	-0.008	-0.008*
	(0.006)	(0.007)	(0.004)	(0.007)	(0.004)
T^{15} (α_2)	0.003	-0.007	0.006*	0.008	0.009**
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)
Control Mean	0.40	0.58	0.25	0.37	0.69
Observations	12895	12895	12895	12895	12895

Unit of observation is the polling station in 2016. Controls include turnout, incumbent turnout, opposition turnout, and distance, all measured in 2011.

All specifications are estimated using OLS (Equation 4) within a pooled bandwidth of +/-200 registered voters. Standard errors clustered at the 2013/2015 polling place-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 6: Reduced form effects of the 2013 reorganization by the distance of previously unregistered voters
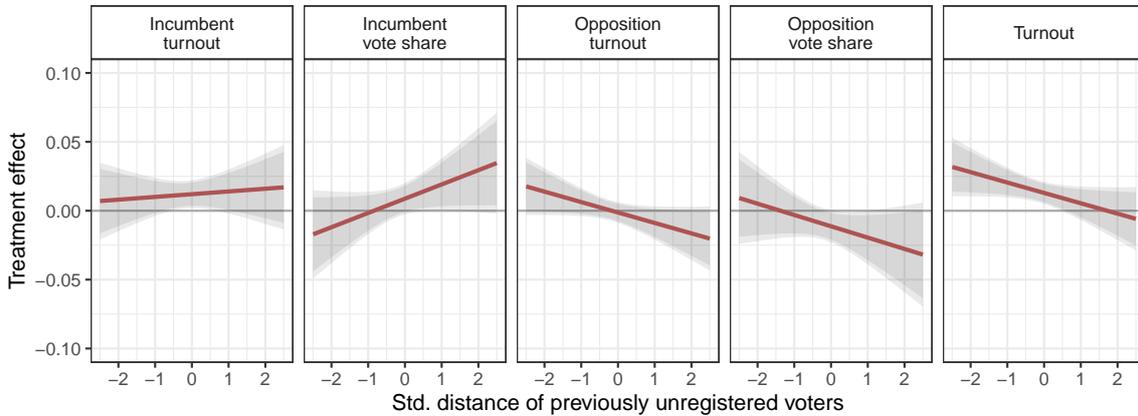
90% and 95% confidence intervals shaded in gray.

In the Appendix, we demonstrate the robustness of each set of estimates to (1) different bandwidths (Figures A6, A7, A8), (2) different samples (Tables A5 and A7), and (3) different specifications (Tables A8 and A9). All results are consistent with our baseline estimates.

The polling places that resulted from the 2013 and 2015 reorganizations thus appear to have an ambiguous overall effect on the level of electoral support enjoyed by the incumbent. While the reduction in the number of registered voters per polling station improved incumbent turnout, the equivalent reductions in distance to vote caused primarily by the 2015 reorganization did not help incumbent turnout or vote share.

However, a back-of-the-envelope calculation suggests that the benefits for the incumbent of the 2013 and 2015 reorganizations in Uganda outweigh the costs. First, we identify every case of a polling place split due to the 2013 reorganization ($n = 6,780$) or the 2015 reorganization ($n = 2,222$). Second, we apply the relevant estimated local average treatment effects from Table 3 to each split polling place and multiply by the total number of registered voters to compute how many additional votes the incumbent received as a result of the splits. Third, we sum this quantity across all polling stations in 2016. This exercise suggests that around 53,000 votes are attributable to the splitting of polling places, which corresponds to around 0.5% of the incumbent's overall vote share. While insufficient to have any decisive effect on aggregate electoral outcomes, the calculation demonstrates how seemingly impartial administrative rules can have meaningful electoral consequences, and thus shape incumbents' incentives to influence electoral administration.

8 Discussion

To reiterate, we find strong evidence that the incumbent NRM party in Uganda—which is in a position to shape the country's electoral infrastructure—faces a trade-off in its evaluation of

whether to increase the number of polling places. Doing so reduces the distance voters have to travel to cast their vote, which levels the playing field between incumbent and challenger parties. When voters have to travel long distances, the resource-rich incumbent is better able than the opposition to control the electorate by selectively mobilizing supporters. However, this effect is dominated by the fact that increasing the number of polling places simultaneously reduces the number of registered voters per polling place. This favors the incumbent because its improved information about local levels of support incentivizes both brokers and groups of voters.

Even though these results draw on an empirical context specific to Uganda, the conditions necessary for incumbents to influence electoral administration in pursuit of electoral benefits are not unique. To assess how the basic forces we identify travel beyond our context, we collected an original dataset of the average number of registered voters per polling station in use for presidential elections in sub-Saharan African countries since 1990.[19] Importantly, there is significant variation across countries in the average number of registered voters assigned to a polling station. The data sources and collection is described in Appendix A.3.

We then consider how this relates to incumbent influence. First, we consider how properties of each country's regime type affect the degree of independence of its electoral authorities. In the top row of Figure 7 we categorize countries as a democracy or autocracy (top-left panel) and as having strong or weak presidentialism (top-right panel). We find that autocratic countries, and countries with strong presidentialism, have notably less independent electoral authorities. This is the case across three different measures of EC independence, with the most conservative estimate suggesting a 10pp difference. While many incumbents are likely to face incentives to control the expansion of polling places, then, we should expect such incentives to be acted upon in countries with less independent electoral authorities—that is, in autocracies and strong

---

[19]This comprises 45 unique countries, and 283 country-election observations. The dataset contains 74% of all presidential elections since 1990 and 91% of all presidential elections since 2000.

presidentialism systems.
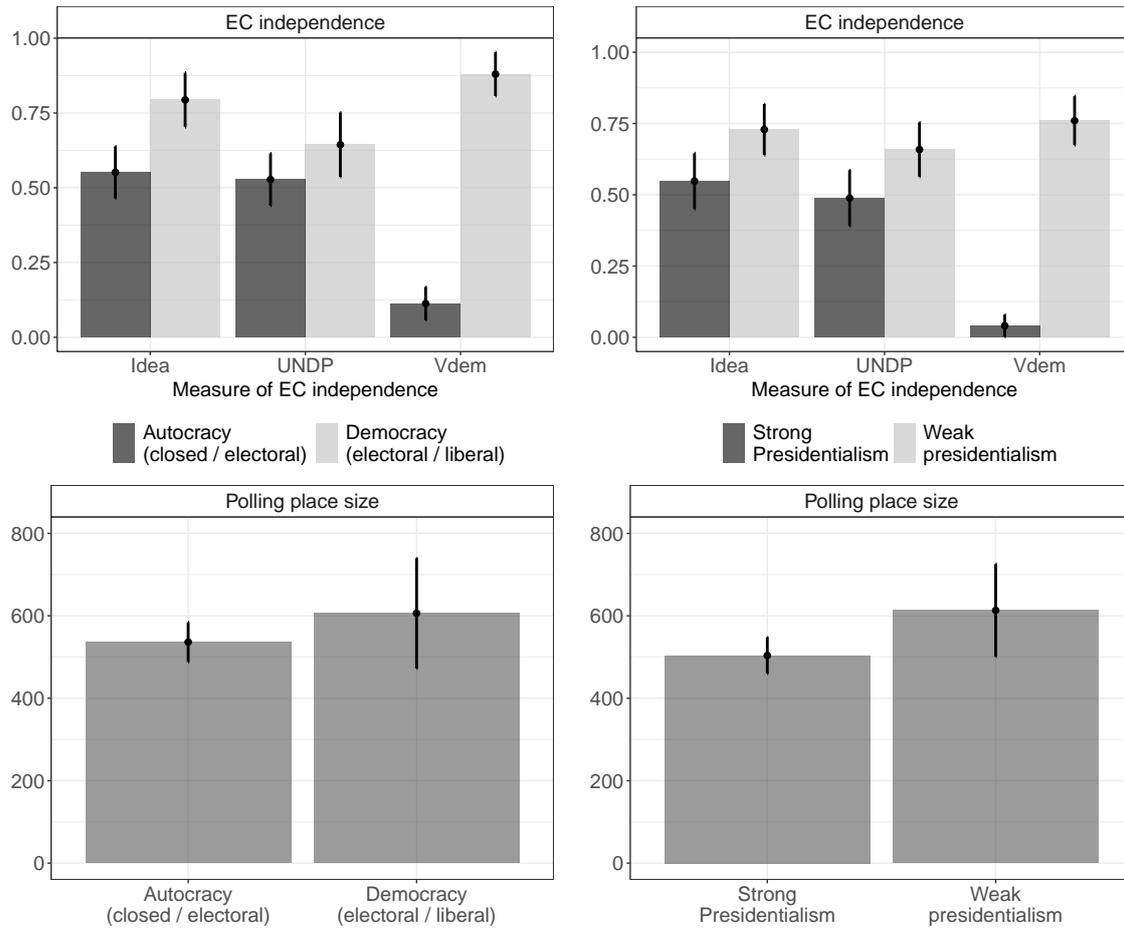
Figure 7: Cross-national comparisons

Next, we consider how the average number of registered voters per polling place varies by regime type depicted in the bottom row of Figure 7. The figures illustrate how countries where the incumbent has considerable influence over electoral authorities have fewer registered voters per polling place than countries where the incumbent has less control over the electoral authorities. In both cases, the difference in size carries practical importance, as countries where incumbents have less control over election authorities have about 100 more registered voters per polling place, or 15% of the total polling place size, than countries where incumbents have more control.

If the expansion of polling places were driven only by sincere democratic motives, then the

fact that polling places are smaller in autocratic regimes is counterintuitive. These cross-national comparisons therefore suggest that the strategic design of polling places is not exclusively a concern in Uganda. Instead, it is a risk in all countries where the incumbent exercises substantial control over the electoral administration. As more than 50% of sub-Saharan Africa countries are characterized by authoritarianism or an otherwise powerful executive, either subtle or overt interventions into the design of voting infrastructure appears to be a regional concern.

9   Conclusion

Relatively little attention has been paid to how incumbents manipulate voting infrastructure in order to maximize their support in contexts where electoral authorities lack autonomy. To address this gap, we empirically investigate the incentives faced by incumbents to manipulate a fundamental part of the election voting infrastructure: the number and location of polling places.

We argue that, confronted with the decision of whether to expand the number of polling places prior to an election, the incumbent faces a trade-off. On the one hand, expanding the number of polling places reduces the number of registered voters at each polling station. This provides the incumbent party's political machine with a higher quality signal about broker performance and collective voting behavior, which has been proven to *increase* the electoral support for the incumbent because the incumbent can use this information to reward individual brokers and groups of voters who support the party. On the other hand, increasing the number of polling places simultaneously reduces the distance traveled by voters to cast ballots. This undercuts the incumbent's control over the electorate, an advantage it holds over opposition parties due to the resources it can leverage to buy turnout, and thus reduces electoral support for the incumbent. Incumbent incentives to introduce new polling places therefore come down to which mechanism

36

has the strongest effect.

We empirically study this trade-off in the context of the 2016 Ugandan presidential elections. Uganda is a typical electoral autocracy, with an electoral playing field heavily skewed towards the incumbent party. Studying the trade off quantitatively is challenging because both the number of registered voters and the distance traveled by voters to the polling place are endogenous. Consequently, to empirically identify the effects using an instrumental variable design, we need at least two sufficiently independent instruments. To surmount this challenge, we use a regression discontinuity design that exploits the repeated application of an administrative rule, implemented in 2013 and 2015, in which polling places exceeding 900 registered voters were split into multiple geographically dispersed polling places. We find that incumbent parties do indeed face a trade-off. Specifically, we find that splitting polling places has a positive effect on incumbent electoral outcomes by reducing the number of registered voters per polling place but a negative effect by reducing the distance voters have to travel to vote. On balance, we find that the reorganizations led to positive effects on incumbent electoral outcomes, suggesting the gains from better information outweigh the costs of reduced control over the electorate.

The insights provided have important implications for democratic consolidation in sub-Saharan Africa and beyond. Encouraged by the recommendations from international organizations, African governments have reorganized their voting infrastructure with the stated aim of heightening political participation. These initiatives are driven by stated expectation that such reorganization empowers citizens by removing barriers to vote, specifically by reducing distances voters have to travel to cast their votes. Consequently, such initiatives are believed to have positive effects on the quality of democracy. Our results imply that such positive effects are limited by a negative effect: an increase in polling places improves the efficiency of political machines. This, in turn, improves the electoral fortunes of incumbent candidates who enjoy a

comparative advantage in turnout buying. Building on these insights, the natural next step would thus be to rethink ways in which to enhance democratic participation without simultaneously facilitating incumbents' use of clientelist voter mobilization strategies.

References

ACFIM (2015). *Who Pays the Piper: Alliance for Campaign Finance Monitoring MP Survey on Commercialization of Politics*. Tech. rep. January.

Afrobarometer (2012). *Afrobarometer Data, Uganda, Round 5*. Tech. rep.

— (2015). *Afrobarometer Data, Uganda, Round 6*. Tech. rep.

— (2018). *Afrobarometer Data, Uganda, Round 7, 2018*. Tech. rep.

Andrews, Isaiah, James H Stock, and Liyang Sun (2019). "Weak Instruments in Instrumental Variables Regression: Theory and Practice". In: *Annual Review of Economics* 11, pp. 727–753.

Birch, Sarah (2012). *Electoral Malpractice*. New York: Oxford University Press.

Blattman, Chris, Horacio Larreguy, Benjamin Marx, and Otis Reid (2020). "Eat Widely, Vote Wisely? Lessons from a Campaign Against Vote Buying in Uganda". In: *Working Paper*.

Bowles, Jeremy, Horacio Larreguy, and Shelley Liu (2019). "How Weakly Institutionalized Parties Monitor Brokers in Developing Democracies: Evidence from Post-conflict Liberia". In: *American Journal of Political Science, forthcoming*.

Brady, Henry E and John E McNulty (2011). "Turning out to vote: The costs of finding and getting to the polling place". In: *American Political Science Review* 105.1, pp. 115–134.

Cantoni, Enrico (2019). "A Precinct Too Far: Turnout and Voting Costs". In: *American Economic Journal: Applied Economics, forthcoming*.

Cantú, Francisco (2014). "Identifying Electoral Irregularities in Mexican Local Elections". In: *American Journal of Political Science* 58.4, pp. 936–951.

— (2019). "The Fingerprints of Fraud: Evidence from Mexico's 1988 Presidential Election". In: *American Political Science Review* 113.3, pp. 710–726.

Chernykh, Svitlana and Milan W Svolik (2015). "Third-Party Actors and the Success of Democracy: How Electoral Commissions, Courts, and Observers Shape Incentives for Electoral Manipulation and Post-election Protests". In: *The Journal of Politics* 77.2, pp. 407–420.

Collord, Michaela (2016). "From the electoral battleground to the parliamentary arena: understanding intra-elite bargaining in Uganda's National Resistance Movement". In: *Journal of Eastern African Studies* 10.4, pp. 639–659.

Coppedge, Michael et al. (2019). "V-Dem Codebook v9 - Varieties of Democracy (V-Dem) Project." In:

De Kadt, Daniel (2019). "Bringing the Polls to the People: How Electoral Access Encourages Turnout but Shapes Political Inequality". In: *Working Paper*.

EC (2019). "The National Voters Register - Guidelines for Verification of Voters at Village Level and Identification of Persons With Disabilities In the National Voter Register". In: *The Electoral Commission*.

EOM (2016). "Uganda Presidential, Parliamentary and Local Council Elections". In: *Election Observation Mission, European Union*.

Finan, Frederico and Laura Schechter (2012). "Vote-Buying and Reciprocity". In: *Econometrica* 80.2, pp. 863–881.

Gans-Morse, Jordan, Sebastian Mazzuca, and Simeon Nichter (2014). "Varieties of clientelism: Machine politics during elections". In: *American Journal of Political Science* 58.2, pp. 415–432.

Gingerich, Daniel W. (2020). "Buying Power: Electoral Strategy before the Secret Vote". In: *American Politcal Science Review, forthcoming*.

Gottlieb, Jessica and Horacio Larreguy (2020). "An Informational Theory of Electoral Targeting in Young Clientelistic Democracies: Evidence from Senegal". In: *Quarterly Journal of Political Science* 15.1, pp. 73–104.

Grossman, Guy and Kristin Michelitch (2018). "Information Dissemination, Competitive Pressure, and Politician Performance between Elections: A Field Experiment in Uganda". In: *The American Political Science Review* 112.2, pp. 280–301.

Hariri, Jacob Gerner, Mogens Justesen, and Anders Woller (2019). "The Cost of Voting and the Cost of Votes". In: *Presented at the 2018 Annual Meeting of the American Political Science Association*.

Harris, J Andrew (2020). "Election Administration, Resource Allocation, and Turnout: Evidence from Kenya". In: *Comparative Political Studies, forthcoming*.

Harris, J. Andrew and Peter van der Windt (2019). "Electoral Administration in Fledgling Democracies: Experimental Evidence from Kenya". In: *Journal of Politics, forthcoming*.

Haspel, Moshe and H Gibbs Knotts (2005). "Location, location, location: Precinct placement and the costs of voting". In: *The Journal of Politics* 67.2, pp. 560–573.

Hidalgo, F Daniel and Simeon Nichter (2016). "Voter buying: Shaping the electorate through clientelism". In: *American Journal of Political Science* 60.2, pp. 436–455.

Jung, Danielle F and James Long (2018). "Social Sanctioning, Trust, and Voter Turnout in Emerging Democracies". In: *Working Paper*.

Kitschelt, Herbert and Daniel M Kselman (2013). "Economic Development, Democratic Experience, and Political Parties' Linkage Strategies". In: *Comparative Political Studies* 46.11, pp. 1453–1484.

Kitschelt, Herbert and Steven I. Wilkinson (2007). "Citizen-Politician Linkages: An Introduction". In: *Patrons, Clients, and Policies: Patterns of Democratic Accountability and Political*

Competition. Ed. by Herbert Kitschelt and Steven I. Wilkinson. Cambridge University Press, pp. 1–47.

Kleibergen, Frank and Richard Paap (2006). "Generalized Reduced Rank Tests Using the Singular Value Decomposition". In: *Journal of econometrics* 133.1, pp. 97–126.

Larreguy, Horacio (2013). "Monitoring Political Brokers: Evidence from Clientelistic Networks in Mexico". In: *Working paper.*

Larreguy, Horacio, John Marshall, and Pablo Querubin (2016). "Parties, Brokers and Voter Mobilization: How Turnout Buying Depends Upon the Party's Capacity to Monitor Brokers". In: *American Political Science Review* 100.1, pp. 160–179.

Newsweek (2016). *Uganda 2016: Polls Open Amid Delays.* Tech. rep.

Nichter, Simeon (2008). "Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot". In: *American Political Science Review* 102.1, pp. 19–31.

Platas, Melina R and Pia J Raffler (2020). "Closing the Gap: Information and Mass Support in a Dominant Party Regime". In: *Journal of Politics, forthcoming.*

Rosenzweig, Leah R (2019). "Social Voting in Semi-Authoritarian Systems". In: *Working paper.*

Rueda, Miguel R (2017). "Small Aggregates, Big Manipulation: Vote Buying Enforcement and Collective Monitoring". In: *American Journal of Political Science* 61.1, pp. 163–177.

Sanderson, Eleanor and Frank Windmeijer (2016). "A Weak Instrument F-test in Linear IV Models with Multiple Endogenous Variables". In: *Journal of Econometrics* 190.2, pp. 212–221.

Simpser, Alberto (2013). *Why Governments and Parties Manipulate Elections: Theory, Practice, and Implications.* New York: Cambridge University Press.

Staiger, Douglas and James H Stock (1997). "Instrumental Variables Regression with Weak Instruments". In: *Econometrica* 65.3, pp. 557–586.

Stock, James H and Motohiro Yogo (2005). "Testing for Weak Instruments in Linear IV Regression". In: *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge University Press, pp. 80–108.

Stokes, Susan C. (2005). "Perverse Accountability: A Formal Model of Machine Politics with Evidence from Argentina". In: *American Political Science Review* 99.3, pp. 315–325.

Szwarcberg, Mariela (2014). "Political parties and rallies in Latin America". In: *Party Politics* 20.3, pp. 456–466.

The Guardian (2016). *Ugandan opposition candidate briefly detained as voting delays criticised*. Tech. rep.

Tripp, Aili Mari (2010). *Museveni's Uganda: paradoxes of power in a hybrid regime*. Vol. 6. Lynne Rienner Publishers Boulder, CO.

Wong, Stan Hok-Wui (2019). "Gerrymandering in Electoral Autocracies: Evidence from Hong Kong". In: *British Journal of Political Science* 49.2, pp. 579–610.

Online Appendix

A Supplementary information

The following section describes our approach to matching polling places across datasets (A.1) as well as how we geocode polling places and calculate average weighted distances (A.2). The final section describes the data processing behind the cross-national data used in the paper (A.3).

A.1 Matching polling places across elections

Sources. For matching polling places over time we rely on five administrative datasets. The two central datasets are made up of lists of the official polling stations in use for the 2011 ($n = $ 23,968) and 2016 elections ($n = $ 28,010). In addition, we make use of the full voter registers for 2011 ($n = $ 13.7 million voters) and 2016 ($n = $ 15.3 million voters) elections. For the 2015 reorganization, we use the number of registered voters for villages in each polling station after the first reorganization but before the second reorganization (N = 193,000 polling station-village observations).

Aggregate polling station to polling place level. In the first step we aggregate polling stations to polling places. This was done by removing any extension after the polling place name (often the venue, e.g. "primary school x", "x trade center") indicating the specific station, including "[A-D]"; "[E-Z]", "A";"B", "I";"II", "1";"2". Doing this we find 21,593 unique polling places in 2011, 23,730 unique polling places in 2015, and 24,367 unique polling places in 2016.

Harmonize administrative unit names across years. When matching polling places over time, we initially match a string containing the district name, county name, subcounty name, parish name and polling place name. Prior to matching polling places, the second step therefore entails harmonizing administrative unit names across elections. In order to identify the same admin-

istrative unit over time, we rely on a similar approach to the one sketched in Figure A1, and elaborated in the next section. However, instead of using voters and villages, we rely on lower-level administrative units, including subcounties, parishes, and villages. In addition, a research assistant provided invaluable assistance with manual coding.

Match polling places across datasets. A universal polling place numerical identification code exists for polling places in Uganda, but this code proved unreliable when matching polling places across elections. When matching polling places across datasets, we therefore rely on a combination of string matching, as well as matching lower-level units across time. We exerted substantial effort in matching very high percentages across elections, as the polling places hardest to identify were likely our study's central observations: polling places split between elections. Of the polling places within our discontinuity sample, we are able to match 4,937 polling places in 2013 to polling stations in 2016 (failing to match 241) and match 4,009 polling places in 2015 to polling stations in 2016 (failing to match 89).

a) String matching entails matching polling places across datasets when names are not identical, but nonetheless describe the same polling station. This can either be due to typos, or minor variations in name. The string matching procedure includes twenty different regular expressions, including removing all whitespace, removing all punctuation, deleting venue identification (e.g. "primary school"), and splitting up long names. The final match is identified as the polling place in the matching dataset that matches the most times across the different regular expressions. This is implemented either for the full identification string or by leaving out one administrative unit at a time, starting from parish and moving towards district. In instances where this does not lead to a match, two administrative unit names are excluded from the identification string.

The string matching proved particularly useful, as it was used not only to match the 2016

polling places to the pre-2013 polling places (95.9% match) and 2015 polling places (97.8% match), but also the 2016 polling station results (99.9% match), and the geocoded polling stations (90.0% match). Finally, the procedure was used to match pre-2013 polling stations to 2011 electoral results (99.3% match).

Figure A1: Illustrating the two polling place matching approaches
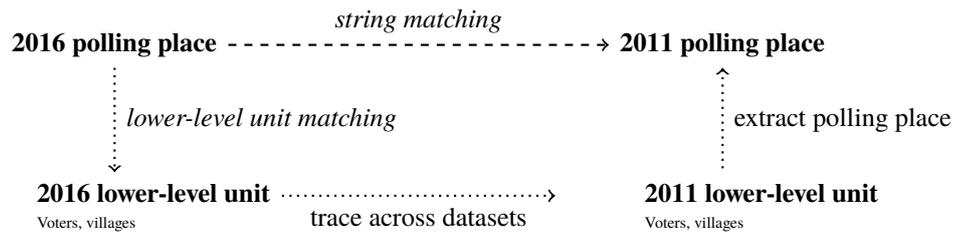
b) Matching using lower-level units was needed in cases where polling places split between elections. The naming convention deems that such new polling places take the name of their new location, and for this reason string matching will not identify matches. When matching polling places, we specifically rely on the corresponding villages and voters. We are able to match voters across the 2011 and 2016 voter registers, and subsequently extract the 2011 polling place name. A match is defined as the polling place in 2011 with the largest share of voters from the 2016 polling place, one requirement being that at least 10% percent of the 2016 voters are identified at the 2011 polling place. We apply the same approach for corresponding villages, which are also available in the voter registers and in 2015. The matching procedure was also applied to match administrative units over time. An example is to match subcounties over time by looking at the share of parishes, polling places or villages matching across datasets.

A.2 Calculating distances to vote

The following section describes how we 1) geocode polling places and 2) calculate distances to the respective villages.

Sources. A central source of data for this exercise is a data set of polling station coordinates made available to us by the Ugandan Electoral Commission. Specifically, we received two datasets, one containing the locations of 25,541 polling stations, while the other contained 12,101 polling station locations. Importantly, the datasets 1) have a number of overlaps (a rough estimate suggests 5,000+ overlaps) 2) are not exhaustive. For the latter reason, we additionally geocode polling stations ourselves. To geocode the 2016 polling places we rely on the 2016 parish demarcations (N = 7,435), administrative data on common polling place locations, including schools, trade centers, health centers, religious institutions, and local administrative headquarters (N = 85,334), containing information on different names of just short of 81,000 administrative units. In order to match administrative unit names in the polling station data to these shapefiles, we rely on the National Geography File (N = 80,866). The file contains the identification codes for the shapefiles containing parishes, subcounties, counties and districts acting as search areas (see below). Finally, we use the 2014 village census (N = 74,277). The latter comes in a polygon format which we turn into coordinates by calculating the centroid.

Geocoding polling stations. First, we match the polling station data to the National Geography File (version 8.0, Oct 28 2016). The National Geography File acts as a "dictionary" file, translating just short of 81,000 administrative unit names and codes between different official datasets. We match the 2016 polling station data to the National Geography File using the share of lower level units, as sketched in Figure A1, with the requirement that the match has to contain at least a third of the lower-level units.

In order to geolocate the polling places, we limit the pool of locations to venues within the given polling station's respective search area by excluding all locations outside the search area using map overlaying. If no match is found within the parish, the search area is expanded to, first, subcounty, then county and, finally, district. Within the search area, we use the string matching

A4

function explained above to identify the polling station in the geocoded data. 23,560 of polling stations are geocoded to locations within their respective parish, 1,782 within the subcounty, 602 within the county and 221 within the district.

Calculating average distance from villages to polling stations. Having successfully geocoded the polling station within the search area, we next identify the villages registered to the polling station in question. Again, we exclude all villages outside the search area using map overlaying. By excluding all villages not located in immediate vicinity to the polling station, we can loosen the string matching requirements substantially. In addition to the search area, we include villages within a buffer of 10 kilometers around the polling station location, in order to not exclude relevant villages which, for one reason or another, are located on the "wrong" side of, say, the parish border, but otherwise in close vicinity to the polling station. It still applies that if a village with a sufficiently similar name is located closer to the polling station, this village is selected.

When we have located the relevant villages, we calculate distances from the villages to the identified polling station. In cases where several polling stations were identified as possible matches, we select the polling station with the shortest average distance to the villages. Polling station-village combinations farther than 20 kilometers are not considered. Doing this, we create a data frame at the village-polling station level, which we aggregate to the polling station unit by averaging all village distances for each polling station, weighted by the number of registered voters at the specific village.[20]

Overall, we identify 84% of the 145,000 village-polling station observations and we match 80% of the villages. For 93.6% of the 2016 polling stations do we identify at least one village. In cases where we identify at least one village, we identify 91 percent of villages on average across polling stations, while we for 3% of polling stations identify less than 50% of villages.

---

[20]For example, if a polling place contains 9 registered voters from village A, which is 0.5 km away, and 3 voters from village B, which is 1 km away, the distance measure is $(\frac{9}{12} \cdot 0.5) + (\frac{3}{12} \cdot 1) = 0.625$ km.

Villages vary in importance, conditioned on how many of its residents vote at the given polling station. Overall, we identify 83.8% of all 2016 registered voters across the 28,010 polling stations. Excluding polling stations for which we do not locate any villages, we identify 87.7% of all voters on average. We identify less than 50% of voters in 12.3% of polling stations. For 7.4% of 2016 polling stations distance is missing. Table A1 shows a comparison of the match between the 2011 and 2015 polling stations with the 2016 polling stations.

Table A1: Matching polling stations and villages

	Matches		Quality of matches		
	Overall	Villages	Villages matched	Voters matched	Less than 50% of villages
	non-missing	matched	per PS* (mean)	per PS* (mean)	matched per PS*
2016	92.6%	84.0%	91.2%	83.8%	3%
2015	87.3%	85.7%	87.1%	71.8%	4%
2011	83.4%	74.1%	82.0%	44.7%	6%

*Polling stations for which at least one village is matched.

A.3 Cross-national data

Sources. For the cross-national data we also rely on a range of sources. For data on elections held in Sub-Saharan Africa, including information on the numbrer of registered voters, we rely on International Idea's database available at `https://www.idea.int/data-tools/continent-view/Africa/40`. We also make use of Idea's categorization of models of electoral management available at the same webpage. Information on regime type and the type of presidentialism (weak or strong, respectively) is found in the Varieties of Democracy dataset v9 (Coppedge 2019). Specifically, we make use of the regimes of the world variable "v2x_regime" and the presidentialism index "v2xnp_pres". We furthermore make use of the variable "v2elembaut_ord" asking if the election management body has autonomy from government to apply election laws and administrative rules impartially in national elections. To measure the impartiality of the

electoral commission we also rely on the UNDP's report on Principles for Independent and Sustainable Electoral Management (see `https://www.ec-undp-electoralassistance.org/`). Information on the number of polling places was collected by the authors, relying on sources including reports published by e.g. the Carter Center, the European Union's Election Observation Missions, as well as international and local media.

Data processing. Two of the central variables, regime and presidentialism, were recoded to binary variables. v2x_regime includes four categories: "closed autocracy", "electoral autocracy", "electoral democracy", and "liberal democracy", which we recoded to "autocracy" or "democracy". v2xnp_pres is an index of the strength of presidentialism, which we recoded to a binary variable depending on whether the observation was above ("strong") or below ("weak") the median. Polling place size was created by dividing the number of registered voters with the number of polling places in use for the election. We study observations from 1995 onwards. The independence of electoral commissions was coded as a binary variable scoring "1" if the electoral commission is independent from the government and membership is expert based, and "0" if the electoral commission is run by the government and membership is partisan. In creating the outcome of interest (either the mean polling place size or mean independence of electoral commission) across regime type and presidentialism, we averaged the outcome of interest within each category. Specifically, every observation was weighted by one over the number of observations which the country in question contributed with. This way, no country affects the overall mean and standard deviation more than others, irrespective of the number of elections included for any particular country.

B Figures
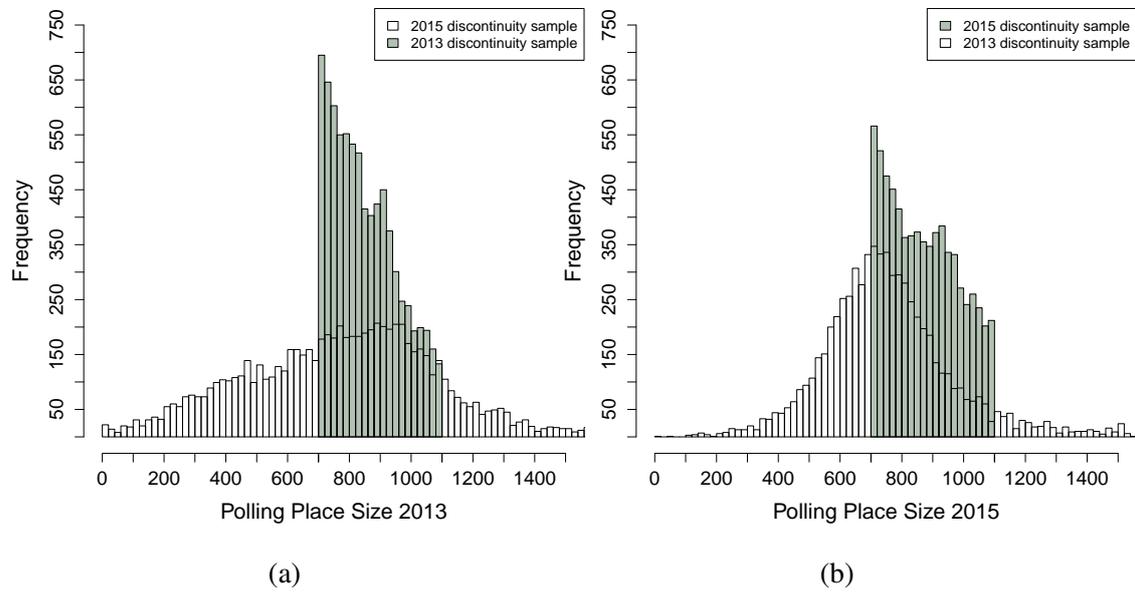
Figure A2: Overlap between discontinuity samples

Overlap is 45.8% in panel (a) and 50.7% in panel (b)
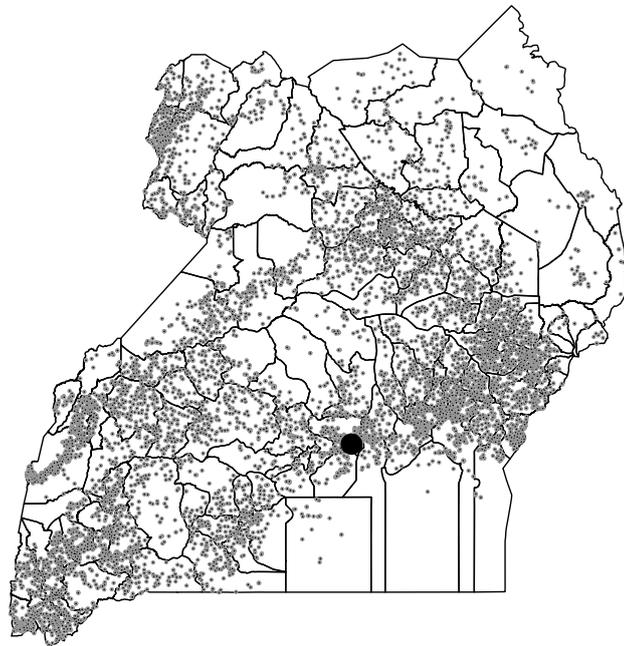
Figure A3: Polling stations included in discontinuity sample

Note Polling stations that appear in our discontinuity sample are shaded in gray. The black dot represents Kampala.

(a)

(b)

(c)
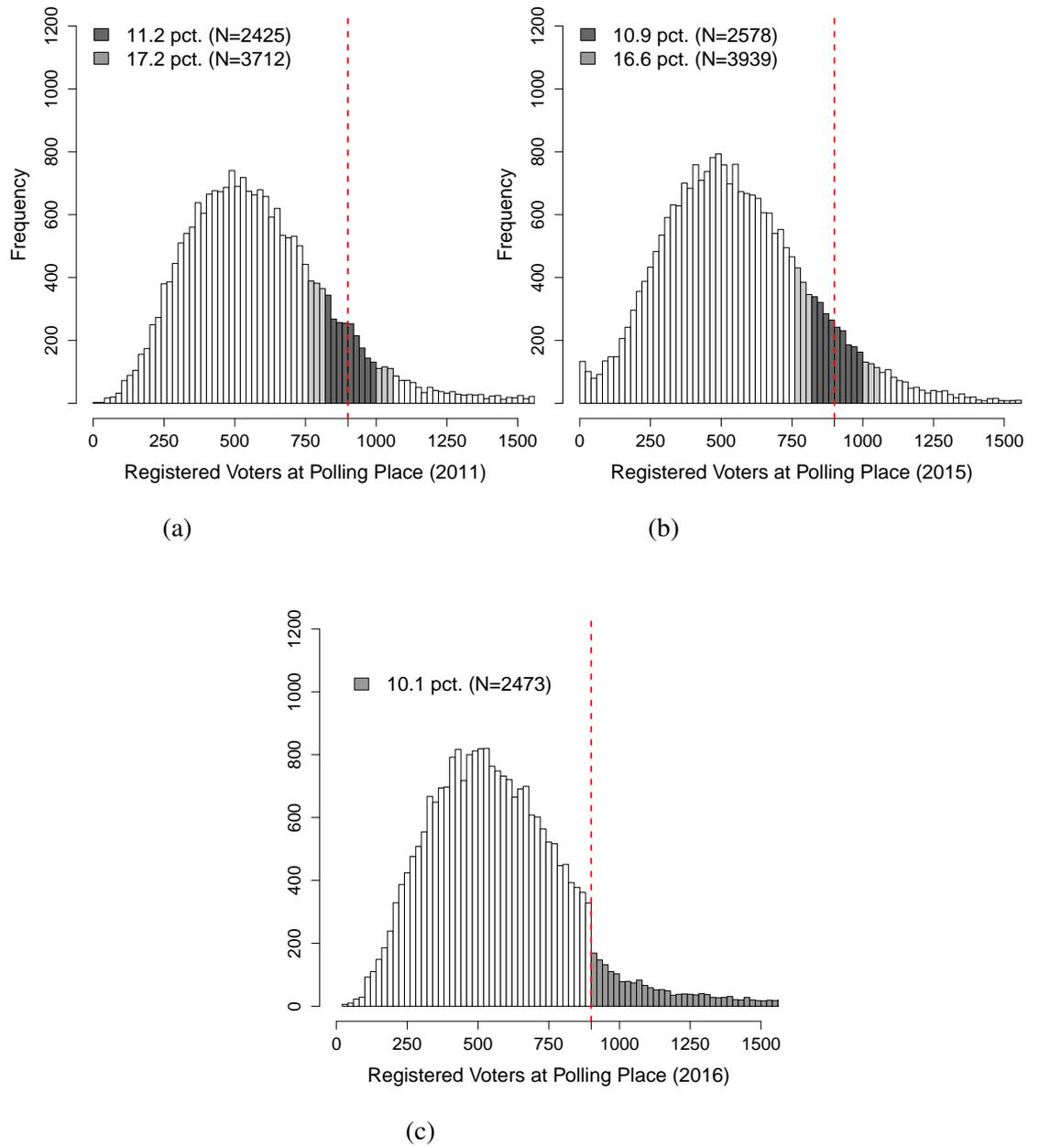
Figure A4: Distributions of polling place sizes over time

Figure illustrates the distribution of registered voters at polling places for the whole sample. From panel (a) and panel (b) it can be observed that a substantial number of observations exist on either side of the 900-voter threshold in both the 2013 and 2015 subsamples, which are used to construct the discontinuity samples. The discontinuous drop observed in 2016 is a direct implication of the administrative rules we leverage in the research design.
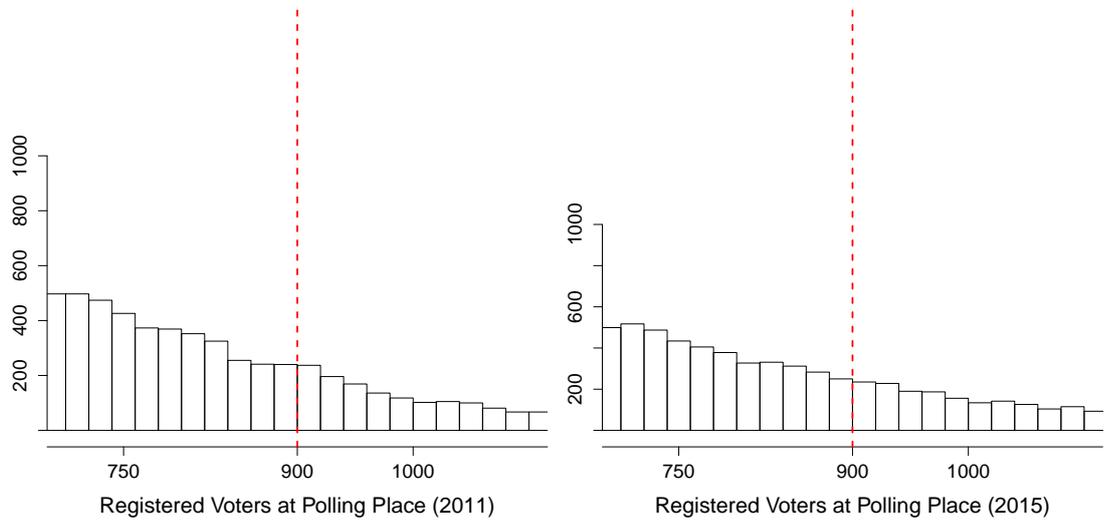
Figure A5: Density of Polling Stations Around Cutoff
Figure indicates that the density of polling places on either side of the threshold in 2013 and 2015 is comparable. By showing that no discontinuity in density exists around the threshold, Figure A5 supports the assumption of no sorting. Balance tests in Table A3 further indicates that no sorting of types seems to take place.

Figure A6: Varying bandwidth for instrumental variable results
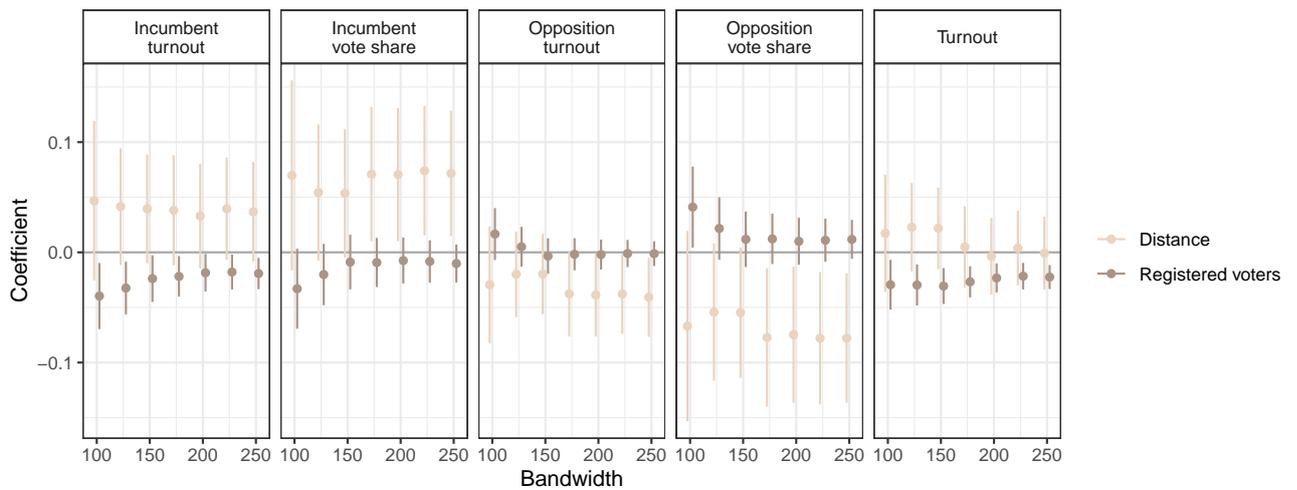
95% confidence intervals plotted. Estimated using Equation (5) with varying pooled bandwidths.

Figure A7: Varying bandwidth for first stage results
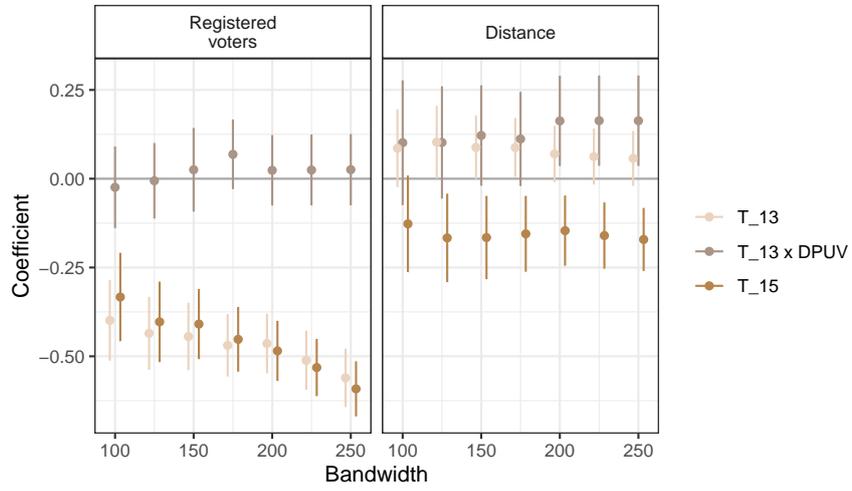
Figure A8: Varying bandwidth for reduced form results
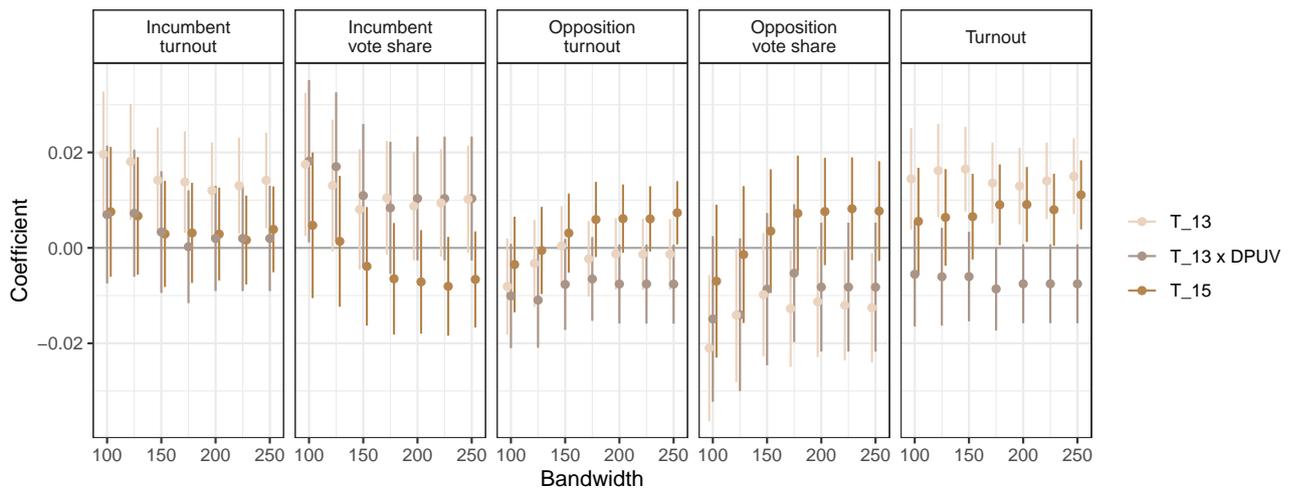
C Tables

Table A2: Descriptive statistics

	Mean	SD	Min	Max
Dependent variables				
Incumbent turnout	0.43	0.16	0.00	1.00
Incumbent vote share	0.60	0.18	0.00	1.00
Opposition turnout	0.24	0.12	0.00	0.73
Opposition vote share	0.35	0.18	0.00	1.00
Turnout	0.71	0.09	0.15	1.00
Independent variables				
Registered voters	539.66	185.01	27.00	1126.00
Distance	1283.82	1584.51	0.00	39177.60
Pre-2016 variables				
Registered voters in PP (2011)	713.17	551.34	17.00	7434.00
Registered voters in PP (2015)	708.35	548.64	4.00	7718.00
Distance (2013)	1456.84	1523.89	0.00	14980.14
Distance (2015)	1327.70	1233.84	0.00	9938.00
Other variables				
PP split (2013)	0.27	0.44	0.00	1.00
PP split (2015)	0.09	0.28	0.00	1.00
Share of voters geolocated	89.75	22.17	0.11	100.00

Descriptive statistics computed at the polling station level in 2016.

Table A3: Balance tests

	Coef. (1)	SE (2)
Incumbent turnout (2011)		
T^{13} (α_1)	-0.03	(0.06)
T^{13} (α_1) × DPUV	0.04	(0.06)
T^{15} (α_2)	-0.05	(0.05)
Incumbent vote share (2011)		
T^{13} (α_1)	0.00	(0.02)
T^{13} (α_1) × DPUV	-0.01	(0.02)
T^{15} (α_2)	0.02	(0.01)
Opposition turnout (2011)		
T^{13} (α_1)	-0.03	(0.06)
T^{13} (α_1) × DPUV	-0.10	(0.07)
T^{15} (α_2)	0.12**	(0.06)
Opposition vote share (2011)		
T^{13} (α_1)	0.00	(0.02)
T^{13} (α_1) × DPUV	-0.01	(0.02)
T^{15} (α_2)	0.02	(0.01)
Turnout (2011)		
T^{13} (α_1)	-0.06	(0.06)
T^{13} (α_1) × DPUV	-0.04	(0.06)
T^{15} (α_2)	0.03	(0.06)
Pre-treatment distance		
T^{13} (α_1)	0.04	(0.06)
T^{13} (α_1) × DPUV	0.04	(0.06)
T^{15} (α_2)	0.16***	(0.06)
Invalid votes (2011)		
T^{13} (α_1)	-0.01	(0.06)
T^{13} (α_1) × DPUV	-0.03	(0.06)
T^{15} (α_2)	-0.02	(0.05)
Longitude		
T^{13} (α_1)	-0.00	(0.06)
T^{13} (α_1) × DPUV	-0.04	(0.06)
T^{15} (α_2)	0.04	(0.05)
Latitude		
T^{13} (α_1)	0.13**	(0.07)
T^{13} (α_1) × DPUV	-0.07	(0.07)
T^{15} (α_2)	0.01	(0.05)

Estimates from regressing pre-treatment covariates defined pre-2013 reorganization onto the right hand side of Equation 4.

All specifications are estimated using OLS (Equation 4) within a pooled bandwidth of +/-200 registered voters. Standard errors clustered at the 2013/2015 polling place-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Exclusion tests

	Coef.	SE
	(1)	(2)
Distance to road		
T^{13} (α_1)	0.05	(0.06)
T^{13} (α_1) $\times$ DPUV	0.09	(0.07)
T^{15} (α_2)	0.08	(0.06)
Distance to village		
T^{13} (α_1)	-0.08	(0.07)
T^{13} (α_1) $\times$ DPUV	-0.06	(0.06)
T^{15} (α_2)	-0.10	(0.07)
School		
T^{13} (α_1)	-0.06	(0.06)
T^{13} (α_1) $\times$ DPUV	0.01	(0.06)
T^{15} (α_2)	0.08	(0.05)
Health facility		
T^{13} (α_1)	-0.00	(0.06)
T^{13} (α_1) $\times$ DPUV	-0.01	(0.07)
T^{15} (α_2)	-0.04	(0.06)
Religious building		
T^{13} (α_1)	-0.01	(0.05)
T^{13} (α_1) $\times$ DPUV	0.02	(0.06)
T^{15} (α_2)	0.07	(0.05)
Market		
T^{13} (α_1)	-0.10*	(0.06)
T^{13} (α_1) $\times$ DPUV	-0.05	(0.06)
T^{15} (α_2)	-0.00	(0.06)
Admin building		
T^{13} (α_1)	-0.05	(0.07)
T^{13} (α_1) $\times$ DPUV	0.03	(0.07)
T^{15} (α_2)	0.03	(0.06)

Estimates from regressing covariates defined in 2016 onto the right hand side of Equation 4.

All specifications are estimated using OLS (Equation 4) within a pooled bandwidth of +/- 200 registered voters. Standard errors clustered at the 2013/2015 polling place-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Reduced Form (Robustness)

	Incumbent		Opposition		
	Turnout (1)	Vote share (2)	Turnout (3)	Vote share (4)	Turnout (5)
I. Weighted					
T^{13} (α_1)	0.011**	0.004	0.002	-0.005	0.015***
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)
$T^{13} \times DPUV^{13}$ (α_2)	0.003	0.011*	-0.007*	-0.008	-0.007
	(0.006)	(0.006)	(0.004)	(0.007)	(0.005)
T^{15} (α_2)	0.006	-0.004	0.005	0.004	0.011***
	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)
Control Mean	0.40	0.58	0.25	0.37	0.69
Observations	12895	12895	12895	12895	12895
II. Excluding Kampala					
T^{13} (α_1)	0.010**	0.007	-0.001	-0.009	0.011***
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)
$T^{13} \times DPUV^{13}$ (α_2)	0.004	0.013*	-0.009**	-0.011*	-0.006
	(0.006)	(0.007)	(0.004)	(0.007)	(0.004)
T^{15} (α_2)	0.006	-0.004	0.006	0.004	0.012***
	(0.005)	(0.005)	(0.004)	(0.006)	(0.004)
Control Mean	0.41	0.59	0.25	0.36	0.69
Observations	12354	12354	12354	12354	12354
III. Excluding parishes with 10,000+ voters					
T^{13} (α_1)	0.009*	0.006	-0.001	-0.008	0.011***
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)
$T^{13} \times DPUV^{13}$ (α_2)	0.003	0.012*	-0.009**	-0.010	-0.008*
	(0.006)	(0.007)	(0.004)	(0.007)	(0.004)
T^{15} (α_2)	0.003	-0.007	0.007*	0.008	0.010***
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)
Control Mean	0.41	0.59	0.25	0.36	0.69
Observations	12252	12252	12252	12252	12252
IV. Excluding PS 2 sd or more from a road					
T^{13} (α_1)	0.012**	0.009	-0.001	-0.012*	0.014***
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)
$T^{13} \times DPUV^{13}$ (α_2)	0.003	0.010	-0.006	-0.007	-0.005
	(0.006)	(0.007)	(0.005)	(0.007)	(0.005)
T^{15} (α_2)	0.004	-0.006	0.006	0.007	0.010**
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)
Control Mean	0.40	0.58	0.25	0.38	0.69
Observations	11583	11583	11583	11583	11583

Unit of observation is the polling station in 2016. Panel I weights observations by the number of polling stations in the same polling place in 2013/2015. Panel II excludes Kampala. Panel III excludes polling stations in parishes with more than 10,000 total registered voters. Panel IV excludes polling stations more than 2 standard deviations from a road.

All specifications are estimated using OLS (Equation 4) within a pooled bandwidth of +/- 200 registered voters. Standard errors clustered at the 2013/2015 polling place-level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A6: First stage without PUV inter-
action

	Registered voters	Distance
	(1)	(2)
T^{13} (α_1)	-0.489***	0.070*
	(0.043)	(0.039)
T^{15} (α_2)	-0.478***	-0.147***
	(0.043)	(0.050)
Partial F-statistic	111.59	17.72
Observations	12895	12895

Unit of observation is the polling station in 2016.
Outcome variables: Column (1): Standardized
number of registered voters in polling station;
(2): Standardized distance of voters to polling
station. Controls include turnout, incumbent
turnout, opposition turnout, and distance, all
measured using data from 2011 election. Partial
F-statistic computed following Sanderson and
Windmeijer (2016).

All specifications are estimated using OLS
(Equation 4) within a pooled bandwidth of +/-
200 registered voters but exclude the DPUV
interaction. Standard errors clustered at the
2013/2015 polling place-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Instrumental Variables (Robustness)

	Incumbent		Opposition		
	Turnout (1)	Vote share (2)	Turnout (3)	Vote share (4)	Turnout (5)
I. Weighted					
Registered voters (β_1)	-0.017**	-0.007	-0.002	0.008	-0.021***
	(0.007)	(0.008)	(0.006)	(0.009)	(0.006)
Distance (β_2)	0.022	0.056*	-0.032	-0.054	-0.011
	(0.026)	(0.032)	(0.022)	(0.033)	(0.021)
Weak Instrument F-statistic	5.01	5.01	5.01	5.01	5.01
Control Mean	0.40	0.58	0.25	0.37	0.69
Observations	12895	12895	12895	12895	12895
II. Excluding Kampala					
Registered voters (β_1)	-0.019**	-0.009	-0.000	0.012	-0.023***
	(0.009)	(0.010)	(0.007)	(0.011)	(0.007)
Distance (β_2)	0.025	0.059**	-0.038**	-0.062**	-0.012
	(0.023)	(0.028)	(0.018)	(0.028)	(0.017)
Weak Instrument F-statistic	6.73	6.73	6.73	6.73	6.73
Control Mean	0.41	0.59	0.25	0.36	0.69
Observations	12354	12354	12354	12354	12354
III. Excluding parishes with 10,000+ voters					
Registered voters (β_1)	-0.016*	-0.006	-0.001	0.008	-0.019***
	(0.009)	(0.011)	(0.007)	(0.011)	(0.007)
Distance (β_2)	0.023	0.058**	-0.038**	-0.062**	-0.015
	(0.022)	(0.027)	(0.018)	(0.028)	(0.016)
Weak Instrument F-statistic	7.34	7.34	7.34	7.34	7.34
Control Mean	0.41	0.59	0.25	0.36	0.69
Observations	12252	12252	12252	12252	12252
IV. Excluding PS 2 sd or more from a road					
Registered voters (β_1)	-0.021**	-0.011	-0.002	0.013	-0.026***
	(0.009)	(0.011)	(0.007)	(0.012)	(0.007)
Distance (β_2)	0.031	0.063**	-0.032*	-0.069**	0.002
	(0.024)	(0.030)	(0.019)	(0.031)	(0.018)
Weak Instrument F-statistic	5.49	5.49	5.49	5.49	5.49
Control Mean	0.40	0.58	0.25	0.38	0.69
Observations	11583	11583	11583	11583	11583

Unit of observation is the polling station in 2016. Panel I weights observations by the number of polling stations in the same polling place in 2013/2015. Panel II excludes Kampala. Panel III excludes polling stations in parishes with more than 10,000 total registered voters. Panel IV excludes polling stations more than 2 standard deviations from a road.

All specifications are estimated using 2SLS (Equation 5) within a pooled bandwidth of +/- 200 registered voters. Standard errors clustered at the 2013/2015 polling place-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We provide a full set of specifications where we do not include the interaction $T^{13} \times DPUV^{13}$, and so just use the indicators T^{13} and T^{15} as instruments. In Tables A6, A8, and A9 we accordingly estimate the first stage, instrumental variables, and reduced form specifications. We find that all the coefficients remain similarly estimated when using this alternative set of specifications. Last, to consider the external validity of the results, we show in Table A10 that OLS estimates of the effects of our two endogenous variables on electoral outcomes are strikingly similar to the IV estimates.

Table A8: Instrumental variables without DPUV interaction

	Incumbent		Opposition		
	Turnout (1)	Vote share (2)	Turnout (3)	Vote share (4)	Turnout (5)
Registered voters (β_1)	-0.018**	-0.005	-0.004	0.009	-0.025***
	(0.008)	(0.010)	(0.006)	(0.011)	(0.007)
Distance (β_2)	0.040	0.066**	-0.030	-0.081**	0.021
	(0.029)	(0.034)	(0.021)	(0.036)	(0.022)
Weak Instrument F-statistic	7.98	7.98	7.98	7.98	7.98
Control Mean	0.40	0.58	0.25	0.37	0.69
Observations	12895	12895	12895	12895	12895

Unit of observation is the polling station in 2016. Outcome variables: Column (1): turnout for incumbent; (2) vote share for incumbent; (3) turnout for opposition parties; (4) vote share for opposition parties; (5) overall turnout. Controls include turnout, incumbent turnout, opposition turnout, and distance, all measured in 2011.

All specifications are estimated using 2SLS (Equation 5) within a pooled bandwidth of +/-200 registered voters but excluding the DPUV interaction. Standard errors clustered at the 2013/2015 polling place-level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A9: Reduced form without DPUV interaction

| | Incumbent | | Opposition | | |
| | Turnout | Vote share | Turnout | Vote share | Turnout |
	(1)	(2)	(3)	(4)	(5)
T^{13} (α_1)	0.012**	0.007	-0.000	-0.010*	0.014***
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)
T^{15} (α_2)	0.003	-0.007	0.006*	0.008	0.009**
	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)
Control Mean	0.40	0.58	0.25	0.37	0.69
Observations	12895	12895	12895	12895	12895

Unit of observation is the polling station in 2016. Outcome variables: Column (1): turnout for incumbent; (2) vote share for incumbent; (3) turnout for opposition parties; (4) vote share for opposition parties; (5) overall turnout. Controls include turnout, incumbent turnout, opposition turnout, and distance, all measured in 2011.

All specifications are estimated using OLS (Equation 4) within a pooled bandwidth of +/-200 registered voters but exclude the DPUV interaction. Standard errors clustered at the 2013/2015 polling place-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: OLS estimates

| | Incumbent | | Opposition | | |
| | Turnout | Vote share | Turnout | Vote share | Turnout |
	(1)	(2)	(3)	(4)	(5)
I. Full sample					
Registered voters	-0.019***	-0.007***	-0.002***	0.010***	-0.025***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Distance	0.012***	0.015***	-0.009***	-0.018***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Control Mean	0.43	0.60	0.24	0.36	0.71
Observations	22094	22094	22094	22094	22094
II. Discontinuity sample					
Registered voters	-0.016***	-0.006***	-0.003***	0.008***	-0.021***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Distance	0.011***	0.014***	-0.008***	-0.016***	0.005***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Control Mean	0.40	0.58	0.25	0.37	0.68
Observations	9947	9947	9947	9947	9947

Unit of observation is the polling station in 2016. Outcome variables: Column (1): turnout for incumbent; (2) vote share for incumbent; (3) turnout for opposition parties; (4) vote share for opposition parties; (5) overall turnout. Controls include turnout, incumbent turnout, opposition turnout, and distance, all measured in 2011.

All specifications are estimated using OLS using either the full sample of polling stations (Panel I) or the pooled discontinuity sample (Panel II). Standard errors clustered at the 2013/2015 polling place-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.