How Weakly Institutionalized Parties Monitor Brokers in Developing Democracies: Evidence from Post-conflict Liberia*

Jeremy Bowles

Department of Government
Harvard University
jbowles@g.harvard.edu
1737 Cambridge Street
Cambridge, MA 02138

Horacio Larreguy

Department of Government
Harvard University
hlarreguy@fas.harvard.edu
1737 Cambridge Street, Rm 408
Cambridge, MA 02138

Shelley Liu

Department of Government
Harvard University
shelleyxliu@g.harvard.edu
1737 Cambridge Street
Cambridge, MA 02138

Running Title:

How Weakly Institutionalized Parties Monitor Brokers in Developing Democracies

Keywords:

weakly institutionalized parties; brokers; monitoring; developing democracies; sub-Saharan Africa

*We thank Dominika Koter, Chappell Lawson, John Marshall, and Gwyneth McClendon as well as participants in the Africa Discussion Group at Harvard University, APSA 2017, the Boston Working Group in African Political Economy, the Frontiers of Clientelism Conference at Harvard University, and the Georgetown Political Economy Seminar for helpful comments. We are grateful to the Liberia Institute for Statistics and Geo-Information Services (LISGIS), the National Elections Commission of Liberia (NEC), and the United Nations Mission in Liberia (UNMIL) for generously sharing their data. We also thank representatives from the Congress for Democratic Change (CDC) and the Unity Party (UP), as well as citizens of Liberia, for graciously agreeing to be interviewed for this project. We gratefully acknowledge funding from the Weatherhead Institute for International Affairs at Harvard University.

How Weakly Institutionalized Parties Monitor Brokers in Developing Democracies: Evidence from Post-conflict Liberia

Abstract

Political parties in sub-Saharan Africa's developing democracies are often considered to lack sufficiently sophisticated machines to monitor and incentivize their political brokers. We challenge this view by arguing that the decentralized pyramidal structure of their machines allows them to engage in broker monitoring and incentivizing to mobilize voters, which ultimately improves their electoral performance. This capacity is concentrated (a) among incumbent parties with greater access to resources and (b) where the scope for turnout buying is higher due to the higher costs of voting. Using post-war Liberia to test our argument, we combine rich administrative data with exogenous variation in parties' ability to monitor their brokers. We show that brokers mobilize voters *en masse* to signal effort, that increased monitoring ability improves the incumbent party's electoral performance, and that this is particularly so in precincts in which voters must travel further to vote and thus turnout buying opportunities are greater.

Replication Materials: The data and materials required to verify the computational reproducibility of the results, procedures and analyses in this article are available on the *American Journal of Political Science* Dataverse within the Harvard Dataverse Network, at: https://doi.org/10.7910/DVN/REXZ8H.

Word count: 9,626

1 Introduction

A rapidly expanding literature illustrates the critical role of political intermediaries in brokering electoral outcomes in clientelistic developing democracies. Brokers are contracted to ensure that the citizens under their influence turn out to register and vote for a particular candidate. Much research has focused on the methods brokers use to enforce such behavior (Nichter, 2008; Finan and Schechter, 2012; Gingerich and Medina, 2013; Duarte et al., 2019). Equally importantly, however, political parties must monitor and enforce broker performance (Stokes et al., 2013; Larreguy, Marshall and Querubin, 2016). In what constitutes a classical agency problem, parties must construct noisy signals about brokers' performance and set up signal-contingent incentives to maximize their voter-mobilization efforts and, consequently, their electoral prospects (Larreguy, 2013; Stokes et al., 2013; Szwarcberg, 2014). Improved monitoring capacity increases broker effort, which improves electoral results.

Research on this principal–agent relationship focuses on the use of sophisticated methods to monitor brokers and thus primarily on more institutionalized democracies (Calvo and Murillo, 2004; Magaloni, 2006; Stokes et al., 2013). The literature on sub-Saharan Africa has paid relatively little attention to how parties monitor and enforce broker effort. This gap can be explained by the seemingly convincing argument that young democracies in sub-Saharan Africa are defined by weak parties lacking the resources and infrastructure to monitor brokers' behavior (Bratton, 2008; Vicente and Wantchekon, 2009; Kramon, 2016).

We contend that the sophistication needed to monitor and reward political brokers has been overstated in the literature. Parties in sub-Saharan Africa often do monitor their brokers' citizen mobilization efforts during voter registration and election periods through a decentralized pyramidal party machine. We also argue that incumbent parties are disproportionately more likely to use decentralized machines to monitor and incentivize brokers because they have access to more resources (Kitschelt and Kselman, 2013). Greater monitoring capacity should therefore increase electoral support for the incumbent, especially in areas in which citizens face higher costs of electoral participation, which increases the scope for turnout buying (Nichter, 2008).

We study post-war Liberia, which, like many other African democracies, faces clientelistic politics. Further, as one of the world's youngest democracies and with weakly institutionalized parties, Liberia is a particularly hard case for broker monitoring and enforcement. At the time of our fieldwork, Liberia had only held two presidential elections since the end of its civil war in 2003. During the first, in 2005, both primary contenders were similarly resource-constrained. However, we show qualitative accounts from party representatives and voters suggesting that, by the 2011 presidential election, the incumbent party was able to establish a decentralized pyramidal network of brokers to monitor and enforce their effort. Access to state resources helped the party fund its decentralized machine, allowing it to monitor and incentivize brokers' turnout buying and ultimately to obtain favorable electoral results.

Using rich administrative data, we provide quantitative evidence that these monitoring efforts are systematic and are consistent with the theoretical argument. We first empirically assess whether brokers effectively mobilize voters *en masse* to make their mobilization efforts more visible to the decentralized party machine (Stokes et al., 2013; Szwarcberg, 2014). Second, we analyze whether the incumbent party—but not the main opposition party—experiences greater electoral support when its machine's monitoring capacity is greater (Larreguy, 2013; Larreguy, Marshall and Querubin, 2016). Third, we test whether the effect of increased monitoring on incumbent electoral performance is more pronounced where voters have to travel further to vote and thus the scope for turnout buying is greater.

Our identification strategies exploit features of electoral administration and voter mobilization in Liberia. The assignment of voters to specific polling stations within a precinct depends on their registration time, with voters that register at the same time being more likely to vote in the same polling station.[1] As in many parts of sub-Saharan Africa, brokers—as a way to signal effort—provide mass transportation from the villages in which they operate during the voter registration and election periods. Consequently, as mobilization by brokers increases, voters from a given village should be more likely to end up voting in the same polling station as each other. We refer to the degree to which voters from the same village are concentrated in a polling station as the *homogeneity* of the polling station, which we show is greater in precincts containing smaller polling stations. Moreover, the National Elections Commission (NEC) adds a polling station to a

[1] A *precinct* in Liberia refers to a building which contains one or more *polling stations*.

precinct if there are more than 500 registered voters per polling station in that precinct. Therefore, precincts that are just above that threshold, rather than just below, have additional and smaller polling stations. We leverage this as-if random variation in polling station homogeneity. Since more *homogeneous* stations provide more informative signals of broker effort, additional polling stations thus should randomly increase the ability of the incumbent party's decentralized machine to monitor broker performance and consequently lead to greater electoral support for that party.

Consistent with brokers signaling their effort during voter registration, we find that, when a precinct randomly receives an additional polling station, the last names of voters registered at a polling station in the precinct increase in similarity. We provide evidence that this increase indicates that voters are more likely to have come from the same village. Second, consistent with well-funded parties monitoring electoral outputs to assess brokers' efforts, we show that additional polling stations led to increased support for the incumbent party only in 2011, but not for the main opposition party in either election. These results are consistent with brokers either buying the turnout of incumbent supporters facing high costs of participation, or persuading individuals who faced lower costs. Therefore, to isolate turnout buying as the mechanism driving the results, we show that this gain in electoral support resulting from additional polling stations was present *only* in precincts in which voters had to travel further to vote and thus there were greater opportunities for turnout buying. All results are robust to the choice of specification, bandwidth, and to the addition of controls. We dismiss alternative interpretations of our empirical findings, including the strategic addition of polling stations and increased voting efficiency.

These results support our contention that, even in young democracies with weakly institutionalized parties, incumbent parties can use their control over state resources and their decentralized party machines to monitor and incentivize brokers. While our identification strategy is specific to Liberia, our argument is more widely applicable. For example, recent work provides qualitative evidence on how the decentralized incumbent machines in Uganda and Senegal support vote buying (Blattman et al., 2019; Gottlieb and Larreguy, 2019). Blattman et al. (2019) specifically show that candidates in Uganda use pyramidal structures to monitor broker effort—including rally attendance and placement of marketing materials in villages— as well as turnout and voting outcomes in the polling stations where they operate.

We contribute to the large literature on African voting behavior, which primarily emphasizes the instrumental role of ethnicity (Bates, 1974; Posner, 2005). Recent work points to the role of local ethnic geography in determining vote choice (Ichino and Nathan, 2013) and to the importance of sparse information in sustaining ethnic-based voting (Conroy-Krutz, 2013; Casey, 2015). We highlight the importance of broker monitoring and incentivizing, even by weakly institutionalized parties, to understanding voting behavior in this context. Further, while the literature repeatedly finds evidence of vote buying in Africa, there is little evidence of broker monitoring (Vicente and Wantchekon, 2009; Vicente, 2014; Kramon, 2016), and parties are typically thought to lack large-scale enforcement capacity (Bratton, 2008). Some research looks at how local authority figures affect political mobilization (Baldwin, 2013; Koter, 2013; De Kadt and Larreguy, 2018; Nathan, forthcoming), but the monitoring of their mobilization efforts is rarely considered.

The paper is organized as follows. Section 2 outlines our theoretical argument. Section 3 provides information on Liberian electoral politics. Section 4 introduces qualitative evidence supporting our argument. Section 5 states our hypotheses applied to the Liberian context. Section 6 presents our data sources and identification strategy. Section 7 presents quantitative results supporting our argument and rules out alternative explanations. Section 8 concludes.

2 Theory

The literature highlights the critical role of local political brokers in intermediating vote buying exchanges (Stokes, 2005; Kitschelt and Wilkinson, 2007). However, parties must also enforce broker behavior to ensure the efficient use of resources. This is a complex task, involving both effective broker selection (Brierley and Nathan, 2019) and the use of ideological ties, career concerns, and monitoring to generate incentives (Larreguy, Montiel Olea and Querubin, 2017). The literature, therefore, focuses on sophisticated political machines within institutionalized—and usually dominant—parties, which are characteristic of more developed democracies in Latin America (Calvo and Murillo, 2004; Magaloni, 2006; Stokes et al., 2013). According to Mainwaring (1999), these parties have strong roots in society, ideological consistency, legitimacy as accorded by politicians, and a party organization independent of specific candidates. Supporters and local brokers, maintain ideological allegiance to the party.

Institutionalized dominant parties are usually considered uniquely able to monitor and enforce broker performance on a large scale, particularly because of their strong societal roots and independent party organizations. Their access to stable state resources attracts the most able political brokers through more remunerative prospects (Kitschelt and Kselman, 2013) and limits broker incentives to appropriate party resources earmarked for vote buying (Larreguy, Montiel Olea and Querubin, 2017). These parties' sophisticated political machines aid them in observing brokers' voter mobilization efforts on the ground, which are substantial undertakings requiring a high level of organization.

Political parties' monitoring ability increases with the observability of brokers' efforts. Szwarcberg (2014), for example, highlights how rallies allow local brokers to signal to candidates their capacity to mobilize voters and thus to buy turnout for the election (Nichter, 2008). Zarazaga (2014) also notes that candidates use other signals, such as the number of the candidates' posters and graffiti within the neighborhood, and even how long these posters and graffiti remain undisturbed by rivals, to assess brokers' capacity to mobilize voters. Stokes et al. (2013) similarly note the importance of the observability of the turnout delivered by brokers for broker selection.

Political parties also evaluate individual broker performance using election results, since they can extract signals through the use of electoral outcomes at the polling station level (Larreguy, 2013; Larreguy, Marshall and Querubin, 2016). For example, Larreguy, Marshall and Querubin (2016) highlight how splitting one precinct into multiple polling stations enables parties to differentiate between random voter-level shocks and broker shirking effort, and thus to infer broker effort. Evaluating broker performance through electoral outputs may then depend on a party's ability to collect (a) accurate information about the expected vote share in each locality and (b) detailed electoral data. It may seem that only an institutionalized party with a sophisticated political machine could monitor brokers at numerous locations.

We contend, however, that the level of political sophistication needed to monitor brokers has been overstated by the literature's focus on institutionalized dominant parties. Rather, political parties can use preexisting networks, circumventing the need to establish strongly embedded party structures, and need only develop a simple decentralized pyramidal structure that mirrors the structures of sophisticated political machines. In contrast to institutionalized party structures, this decentralized pyramid structure is more likely to build on personal and informal ties that originated from outside the party, and do not depend on ideological allegiance.

A pyramidal structure allows for different levels of broker monitoring, such that local-level brokers at the base of the pyramid are monitored by intermediate-level brokers, who are in turn monitored by higher-level brokers, who are in turn monitored by the electoral campaign team that reports to the candidate. By delegating the recruiting, incentivizing, and monitoring of lower-level brokers to upper-level brokers, candidates from weakly institutionalized parties turn the large problem of disciplining brokers into multiple simple tasks. While sophisticated large-scale monitoring strategies might be unfeasible for political parties lacking a developed party machine, brokers at each level of the pyramidal structure can simply monitor lower-level brokers' efforts during voter registration and election periods (Koter, 2013; Gottlieb and Larreguy, 2019).

This disaggregated method of monitoring broker performance allows the political party to enforce brokers' voter mobilization efforts. Because lower-level brokers are directly beholden to higher-level brokers, monitoring efforts should be similarly visible to lower-level brokers, who are aware of the ongoing evaluation. Broker monitoring thus ensures increased effort during voter mobilization, as each level of the pyramid is incentivized to achieve greater results from below in order to receive rewards from above. The end result is increased broker enforcement when monitoring ability increases, ensuring favorable electoral returns.

We add to a growing literature on party structures and voter mobilization using non-party networks in contexts with weakly institutionalized parties. For example, Novaes (2015) argues that political parties in Brazil with access to resources can adopt preexisting networks of voters from intermediate-level brokers. Thachil (2011, 2014) similarly document a division of labor between party officials and local mobilizers in some political parties in India, which can outsource the mobilization of poor voters to non-party affiliates. Koter (2013) shows that candidates in Africa subcontract voter mobilization to established local leaders who control voter networks. Likewise, Aspinall (2014) shows how Indonesian candidates mobilize voters through personal, informal networks with a pyramidal structure.

Taken together, this literature illustrates a variety of decentralized mobilization techniques that tap into existing social networks. While this literature addresses how voters can be mobilized by non-party brokers, we show how broker performance can be monitored by weakly institutionalized parties using similar decentralized structures. Moreover, our argument does not rely on coopting organized local organizations. As long as higher-level brokers maintain personal ties to individuals who they can recruit as lower-level brokers and subsequently monitor and incentivize, weakly institutionalized parties can monitor and enforce broker performance on a large scale.

Weakly institutionalized parties, however, require resources to build these decentralized pyramidal structures. Because brokers are not ideologically motivated but are instead contracted agents seeking self-benefit, parties must ensure those benefits to ensure proper monitoring at each level of the pyramid. In developing contexts with resource-constrained parties, we therefore expect the incumbent party to be more likely to be able to build such structures. Incumbent political parties have access to state resources as well as employment opportunities, allowing them to incentivize and monitor brokers. In contrast, opposition parties with fewer resources are far less likely to assemble a decentralized pyramid structure for voter mobilization.

We therefore challenge the predominant view in the African politics literature, which argues that weakly institutionalized parties in Africa lack the capacity to recruit, incentivize, and monitor brokers for voter mobilization (Bratton, 2008; Vicente and Wantchekon, 2009; Kramon, 2016). Building on the work on decentralized party structures discussed above, we contend that candidates from weakly institutionalized parties can establish structures to facilitate broker monitoring particularly if they have access to state resources through incumbency.

3 Context

We focus on the Liberian presidential elections of 2005 and 2011, which were the two elections held between the end of the civil war in 2003 and our fieldwork.[2] These elections were contested between the Unity Party (UP)—the incumbent party from 2005 to 2017—and the Congress for Democratic Change (CDC)—the main opposition party during that period. Political parties traded vote buying accusations each campaign season. Though such complaints were relatively minor

[2] See Appendix A.1 for details on the fieldwork we conducted.

during the 2005 election (Carter Center and National Democratic Institute, 2005), they became more serious in 2011 and culminated in an opposition-led boycott of the runoff election.

By Mainwaring (1999)'s definition, both the UP and the CDC are weakly institutionalized. As nascent parties contesting in the first post-war elections, they were not strongly rooted in society and maintained no ideological stance. Electoral success was highly dependent on specific candidates. After 2005, both parties had begun to build limited support bases. However, party institutionalization remained weak as neither party has been able to develop the local party structures characterizing strongly institutionalized parties. Further, both parties are still defined by their candidates rather than by discernible ideological stances.

3.1 Turnout buying

Accounts suggest high levels of organized turnout buying by the UP and, to a lesser extent, by opposition parties in 2011. In a USAID survey of electoral practices in Liberia, 49% of surveyed citizens responded that "many" or "almost everyone" accepted gifts from parties in exchange for their vote (USAID, 2015). Since a substantial majority of citizens believe in the secrecy of vote choice (Afrobarometer, 2012), political parties use such strategies primarily among likely supporters to encourage turnout. During an interview, a UP representative explained the importance of ensuring that supporters turn out to vote: "You can't forget your stronghold. If you a give a bag of rice to your opposition (...) you have to send two to your stronghold" (Monrovia, July 11, 2016).

Parties also encourage turnout by providing transportation for citizens first to register and subsequently to vote, particularly in rural areas with poor roads. Reflecting its ubiquity, the USAID survey reported that 65% of the respondents heard of voters being transported to polling precincts, and 68% recognized that trucking was "very common" (USAID, 2015). Both parties reported that its brokers aid rural supporters by increasing voter transportation in these areas, although the CDC suggested that its ability to do so is severely limited due to lack of resources (Monrovia, July 6, 2016). However, since the UP came to power in 2005, resources have not been a problem for the party, and it has done its best to facilitate turnout. As a UP representative noted, "there is no law that prohibits me from taking my private car and taking anybody to the ballot box. And if we have all [our] partisans around, then it means this party [will put] all their vehicles around" (Monrovia, July 11, 2016).

3.2 Campaign financing and relative advantage

Parties' ability to buy electoral support is largely determined by their access to funds. This was especially evident in the 2005 election, in which the leading candidates equally struggled to finance their campaigns (Adejumobi, 2006). Due to post-conflict insecurity and destroyed infrastructure, the parties were unable to systematically provide incentives and transportation to mobilize rural citizens. Therefore, civilians and party officials reported only isolated and infrequent instances of vote buying (Harris, 2006). However, after winning the presidency in 2005, the UP was able to greatly increase its capacity.

Observers noted that in the 2011 election, the UP was far better financed due to incumbent control over state resources (Carter Center, 2011; Boas and Utas, 2014). Voters similarly reported an increase in vote buying, with 59% of a representative sample believing that the UP was the main party engaging in vote buying compared to 10% for the CDC (USAID, 2015). As one interviewee recalled, "in 2011, UP did more campaigning than CDC... carrying rice, carrying money. UP has more money than CDC. CDC, throughout Liberia, cannot spend money during elections" (Monrovia, July 13, 2016). Testimony from UP and CDC officials corroborated this difference in the parties' financial and organizational capacities—and, consequently, in vote buying—in 2011. While the UP was able to deploy brokers throughout the country to encourage voter registration and turnout, the CDC continued to struggle with the prohibitive costs of reaching much of rural Liberia. As a CDC official highlighted, "engaging with rural [areas] is capital-intensive. We do not have the resources. We have not ventured to embark on [rural] roads" (Monrovia, July 18, 2016).

3.3 Electoral administration

The 2005 and 2011 presidential elections followed similar timelines. Presidents were elected using a runoff system and, in both years, the first round of elections took place in October and the second in November. Voter registration took place several months before each election, with no permanent voter register existing between these elections. The NEC is responsible for managing electoral administration and was well-staffed and logistically prepared for these tasks (Carter Center, 2011). Moreover, the NEC is widely trusted by voters (USAID, 2015).

The NEC creates precincts, which are the locations that serve as both voter registration and polling places. Voters are registered on a first-come-first-served basis and are assigned sequential ID numbers. During the registration process, citizens are issued a voting card—with their photo and ID number—which must be presented on election day at the location at which they registered in order to vote. In accordance with electoral law, parties receive access to the voter register, including voters' names, ages, ID numbers, precinct locations, and polling station assignments.[3]

After registration, the NEC determines the number of polling stations in each precinct, based on the number of registered voters in the precinct. In 2005, each polling station serviced a maximum of 600 voters, and in 2011, 500 voters. If the number of registered voters in a precinct exceeds this maximum, polling stations are added to the precinct and the total voting population registered to that precinct is divided equally among all polling stations according to their voter ID numbers. Each polling station covers a range of numbers and individuals are assigned to each polling station based on the last three digits of their voter ID number. This process ensures that individuals who register to vote at the same time—and therefore receive consecutive voter ID numbers—are more likely to be assigned to the same polling station. Thus, if a group of people from the same village arrive to register together, they are later likely to vote in the same polling station.

4 Qualitative Evidence

4.1 Decentralized broker structure

Interviews with party officials and voters show how access to state resources allowed the UP to set up a decentralized broker structure for voter mobilization in 2011. UP officials spoke of the pyramidal structure they used to mobilize voters. One official explained that they have brokers everywhere at the base: "Within the town, we have a focus person. In the villages, we have focus persons. In the hamlets, we have focus persons. And they coordinate our party activities there"

[3] While the NEC does not digitize voters' village of residence, residence information is common knowledge for the brokers in charge of specific localities.

(Monrovia, July 10, 2017). There is roughly one low-level broker per each of the villages and hamlets, which are aggregated into approximately 1500 *zones*, each under a zonal leader. Another official indicated that "the district people manage the people in the zones" and that "zonal people work in close consultation with the district people because we follow the chain of command, we have to respect people when you give them responsibility" (Bomi, August 12, 2017).

A UP official explained the need for such a decentralized structure: "You do not know everybody, because that is not possible. In a county, you have to go in[to] the districts, you got to go in[to] the clans, you got to go in[to] the villages. Who takes you there? It's the local campaign managers, the local party leaders, they [are] going to do the work. (...) I like to call it the bottom-top approach. Because if it is the top-bottom approach, you won't get the results. It has to be the bottom-top approach" (Monrovia, July 11, 2016). Reflecting this decentralized structure, an official in the UP headquarters pointed out that they "do not [even] have a central list here of the focus people. [We] only [have] the lists of the county people, of the district people—then they manage the people in the zones and the towns" (Monrovia, July 10, 2017).

4.2 Capacity to incentivize brokers

The UP's access to state resources created strong incentives for people to seek out low-level "focus person" positions within the party and to perform well once hired. When choosing brokers, the party prioritized people who had exhibited some loyalty to the party, due to the risks associated with opportunists (Monrovia, July 11, 2016). Nevertheless, there was still clearly a *quid pro quo* relationship such that the UP was expected to reward brokers who had delivered electoral support. Rather than emphasizing the importance of the party ideology over individual interests as a way to motivate its brokers, a UP official acknowledged that "[I]f the party becomes victorious tomorrow, you have to pay [the broker] back" (Monrovia, July 11, 2016).

Thus, incentives such as monetary rewards and employment opportunities played a key role in encouraging broker effort, especially when such effort was visible to those higher up in the pyramid. This logic is pervasive in Liberia, where citizens pointedly mentioned that those working for the UP had incentives to garner votes for it so as to keep benefiting from affiliation with the incumbent party (Monrovia, July 13, 2016).

4.3 Monitoring brokers' effort

The UP complemented its broker recruitment with close monitoring of their efforts to register and turn out citizens to vote. Such efforts are rendered relatively observable by the brokers' intensive provision of transport in delivering citizens to polling stations. To begin with, as a UP official explained: "We are involved in the voter registration process. We have observers at the polling centers making sure registrations are done." (Monrovia, July 11, 2016). Another UP official explained how they monitored the mobilization efforts of their local brokers: "Your village [is] from the countryside, they come together to the town (...) You have your zonal leader in the precinct and they can see—hey, this village did not turn out" (Monrovia, July 10, 2017).

The brokers at intermediate tiers of the UP pyramidal structure also had complementary strategies to evaluate the lower-tier brokers' work. During focus group discussions, citizens explained that voters must provide their voter registration ID numbers to lower-tier brokers when accepting a gift in exchange for their vote (Monrovia, July 12 and 19, 2016). Lower-tier brokers then turn over these ID numbers to the intermediate-level brokers, who cross-reference them with the official voter register to keep track of the allocation of voters under specific lower-tier brokers to particular polling stations within each precinct. The voter makeup of each polling station and the information about the voters under each of the lower-tier brokers facilitated the monitoring of their performance on election day.

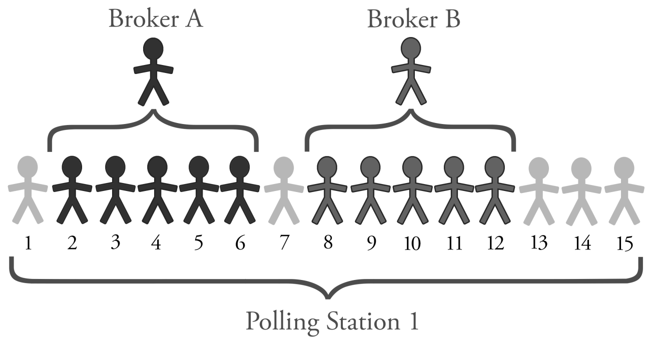
Second, the intermediate-tier brokers assigned polling monitors to every precinct on election day to observe the voting and the tallying process for each polling station. As the polling station results were reported, party officials convened to determine whether each station had met its target so they could reward the lower-tier brokers who had delivered the expected electoral support. As a UP official indicated, they "look at the [election] results of the precinct, to say whether this precinct does worse than the ones around it" (Bomi, August 12, 2017). Another explained, it is important to conduct "post-election analysis (...) to know which of our focus persons are doing [good] work (...) everyone is going to tell you 'I'm the best, I'm the best' " (Monrovia, July 11, 2016). Similarly, party officials then determine whether higher-level brokers at the district level are capable monitoring and incentivizing lower-level brokers to mobilize voters within their areas. Higher-level brokers who are in charge of with poor electoral performance are then replaced the next election (Monrovia, July 11, 2016). Rewards generally include moving up the party ladder or even government jobs.

All these accounts of the UP contrast greatly with accounts of the CDC. Based on our interviews and focus groups, individuals consistently noted that the CDC struggled to deploy brokers, particularly in rural areas (Monrovia, July 18, 2016). Instead, the CDC largely relied on volunteers and, due to financial constraints, was unable to monitor and incentivize its brokers in an organized fashion. It largely relied on party loyalty; its monitoring apparatus was limited to ensuring that its brokers exerted minimal effort. Interviews with party officials indicated that the party did not systematically monitor broker effectiveness using polling station–level election results (Monrovia, July 6, 2016).

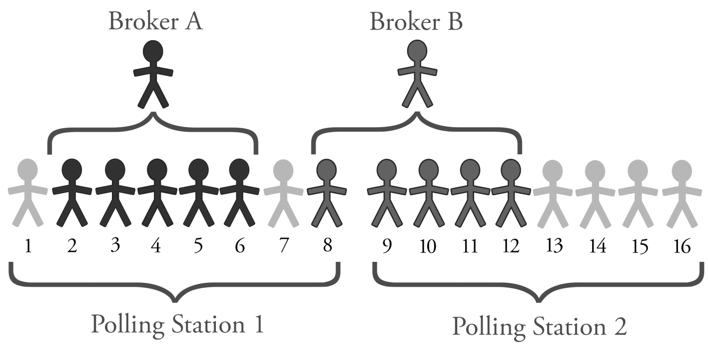
5 Quantitative Hypotheses

The qualitative evidence shows that the UP set up a decentralized pyramidal broker structure for mass voter mobilization in 2011. Exploiting state resources, the UP recruited individuals and monitored their mobilization efforts to ensure that they delivered strong electoral support. In this section, we derive our theoretical argument's empirical implications for the Liberian context to assess whether the behavior described in the qualitative accounts was systematic.

We first test whether low-level UP brokers systematically mobilized voters to signal their mobilization effort to higher-up brokers. We focus on the voter registration period and leverage features of the voter registration process and the splitting of precincts into polling stations, which we detailed in Section 3. Since voters registering at the same time receive consecutive voter ID numbers and each polling station covers a range of consecutive numbers, then if brokers encourage voter registration by providing transportation or otherwise organizing large groups of rural voters to register together, voters who live within a particular broker's area of operation should be more likely to be assigned to the same polling station. Moreover, as we illustrate in Figure 1, when polling stations are added to a precinct, this likelihood should be even greater. In other words, voters mobilized by brokers should be more likely to end up concentrated in a few *homogeneous* polling stations, where registered voters come from fewer villages.



(a) *Heterogeneous* polling station



(b) *Homogeneous* polling stations

Figure 1: The effect of an additional polling station on polling station homogeneity

In the simple example in Figure 1, a maximum of 15 voters can register at a polling station before another polling station is added in the precinct. Most registered voters come from villages A and B, where the corresponding brokers each mobilizes five voters. Under scenario 1a, 15 voters registered in the precinct, which results in one heterogeneous polling station. Under scenario 1b, however, one additional registered voter causes an additional polling station to be added to the precinct and the registered voters are split following their registration order into two more homogeneous polling stations.

As this example illustrates, if brokers effectively encouraged voter registration *en masse* to signal their mobilization efforts, the likelihood that voters registered at a given polling station within a precinct come from the same village—and thus polling station homogeneity—should increase with the number of polling stations in a precinct.

Hypothesis 1 *A polling station within a precinct is more likely to serve voters from a single village—and* *thus a more homogeneous population—when additional polling stations are created for the precinct.*

Next, we test whether parties—in particular, the UP in 2011—systematically used electoral outcomes at the polling-station level to monitor brokers, induce greater mobilization efforts, and ultimately obtain a more favorable electoral performance. We exploit the fact that parties can better monitor brokers when there are more polling stations in a precinct. As additional polling stations are added, the vote totals of a given village become clearer, as do the efforts and resources allocated by the broker mobilizing that village.

A UP official corroborated that additional polling stations capture increased ability to monitor broker performance. In precincts with only one polling station it is hard to determine whether a broker shirked effort since "you have town one ... to town ten. This is a polling center. So if you got, say, 104 votes last election. And this election you got 50 votes... but which one of the towns are you going to say 'you didn't vote for me'?" (Monrovia, July 19, 2016). But, when there are multiple polling stations, "we [observe] by voting card... because it's possible that we all go to one polling station. So with that, we also know which town is voting against" (Monrovia, July 19, 2016).

This increased monitoring capacity should induce broker effort, leading to better electoral outcomes. Thus, if the UP systematically monitored brokers' efforts in 2011 using election results, there should be an increase in electoral support for the UP in precincts with additional polling stations. However, this effect of monitoring on increased support for the party should be absent for both parties in 2005 and for the CDC in 2011 due to lack of resources.

Hypothesis 2 *Precincts to which polling stations have been added should have higher electoral support for the UP in 2011 but not in 2005 and not for the CDC in either year.*

Increased electoral support is consistent with either the turnout buying of individuals facing high costs to participation, or the persuasion of individuals facing lower costs. Our fieldwork suggests that there is greater scope for turnout buying in rural areas, since voters who live further from their polling stations are less likely to turn out to vote and are thus more likely to be subject to targeted turnout buying by brokers. We then expect that the effect of an additional polling station on party electoral support should then be increasing in the distance that voters must travel to vote. As with Hypothesis 2, we expect this relationship to hold only for the UP in 2011.

Hypothesis 3 *The effect of additional polling stations on UP electoral support in 2011 should be larger for precincts in which voters have to travel further to vote.*

6 Empirical Strategy

In this section, we describe our data and demonstrate how discontinuities in the allocation of additional polling stations provide exogenous variation for testing our hypotheses.

6.1 Data

Table A4 presents summary statistics for our key variables. To analyze polling station homogeneity, we use the complete 2014 Liberian voter register, which adds around 5% of newly eligible voters relative to the 2011 register. Throughout the analysis of voting outcomes, we use precinct-level electoral data from the NEC. There were 1,355 voting precincts in the 2005 election and 1,780 in 2011 across Liberia's 15 counties. We exclude precincts from two very urban districts throughout the analysis for both theoretical and empirical reasons explained later. We focus on the first rounds of the 2005 and 2011 presidential elections, given the CDC boycott during the 2011 runoff election. We also use geocoded census data at the locality level from the 2008 National Population and Housing Census and the geocoded location of each polling station from the United Nations Mission in Liberia to construct precinct-level covariates.

6.1.1 Dependent variables

Polling station homogeneity

H1 indicates that adding polling stations in a precinct should increase the extent to which voters from particular villages are concentrated at specific polling stations. Since the NEC does not collect the addresses of registered voters, we construct dependent variables using data on voters' names from the voter register to measure the last name similarity of voters registered at each polling station within every precinct. As Figure 1 illustrates, assuming that last names are more similar *within* villages than *across* villages due to familial ties, additional polling stations in a precinct should increase last name similarity within polling stations in the precinct.

Employing multiple measures of name similarity is important in contexts such as Liberia where literacy levels are low and the possession of formal identity documents is limited, such that spelling differences for otherwise related individuals are likely to occur. Consequently, we generate three measures of last name similarity, which we justify and explain in more detail in Appendix A.2.1. First, for each registered voter v in polling station p in precinct i, we calculate the share of other registered voters in p for which there is an exact match on last names. Second, we calculate the share of other voters for which there is a Soundex match of phonetic similarity with v's last name. Third, we calculate the average Levenshtein distance of v's last name to all the other last names of voters in p. We then average each of these measures to the precinct level. As the homogeneity of polling stations increases, we expect the exact match and Soundex measures to increase, and the Levenshtein distance measure to decrease. In the analysis, we standardize these measures and create a z-score index aggregating them.

We discuss and validate the logic of using name similarity to measure polling station homogeneity in Appendix A.2.2. First, we use census data to show that precincts close to more ethnically homogeneous localities contain significantly more similarly-named voters. Second, we leverage nationally-representative survey data to show that household heads have significantly more similar last names to other household heads in their enumeration area compared to household heads in neighboring enumeration areas. These validation exercises suggest that last name similarity is an appropriate measure of polling station homogeneity.

Party electoral support

H2 predicts that electoral support in 2011 should increase with additional polling stations within a precinct only for the UP. For measures of electoral support for the UP and the CDC in 2005 and 2011, we consider the number of votes for each party divided by the number of registered voters in each precinct. We do not condition party vote on turnout in the baseline specification so that the measure remains independent of potentially endogenous mobilization efforts. Increases in the dependent variable, therefore, result from increases in the overall share of registered voters voting for the party rather than from reductions in support for rival parties.

6.1.2 Independent variables

The key independent variable is whether a precinct received the treatment of an additional polling station, which we detail in Section 6.2 below. We also examine the heterogeneous effects of distance (H3) by constructing a measure of the weighted distance of a precinct's voting population from their polling stations. We provide details in Appendix Section A.3, where we justify the exclusion of two urban districts for which we have to impute a high proportion of observations. However, we also demonstrate the robustness of our central results to the inclusion of those districts. The median distance to precincts was 2.3 kilometers in 2005 and 1.9 kilometers in 2011. For additional covariates, we draw on the 2008 census.

6.2 Identification strategy

Precincts with more polling stations are likely to systematically differ from those with fewer stations, which confounds the relationship between the number of polling stations and voting outcomes. Our identification strategy, therefore, hinges on exogenous variation in additional polling stations within a particular precinct, which are created once the number of registered voters in a particular precinct exceeds a given threshold (600 in 2005 and 500 in 2011). The plots in Figure 2 indicate that the majority of precincts had three or fewer polling stations in both elections. Importantly for identification, additional polling stations are in the same location as existing ones and thus do not affect the distance that voters must travel, which Appendix Table A5 corroborates.

To exploit exogenous variation in additional polling stations, we use a fuzzy regression discontinuity design (RDD) in which we compare the electoral outcomes of precincts just above and just below the threshold for adding new polling stations. We pool the different thresholds (for example, in the 2011 election, 500, 1000, 1500, and so on) and define our forcing variable as the distance—in number of registered voters—from the closest threshold. The forcing variable thus varies between -250 and 250 for the 2011 election and -300 and 300 for the 2005 election. Figure 2

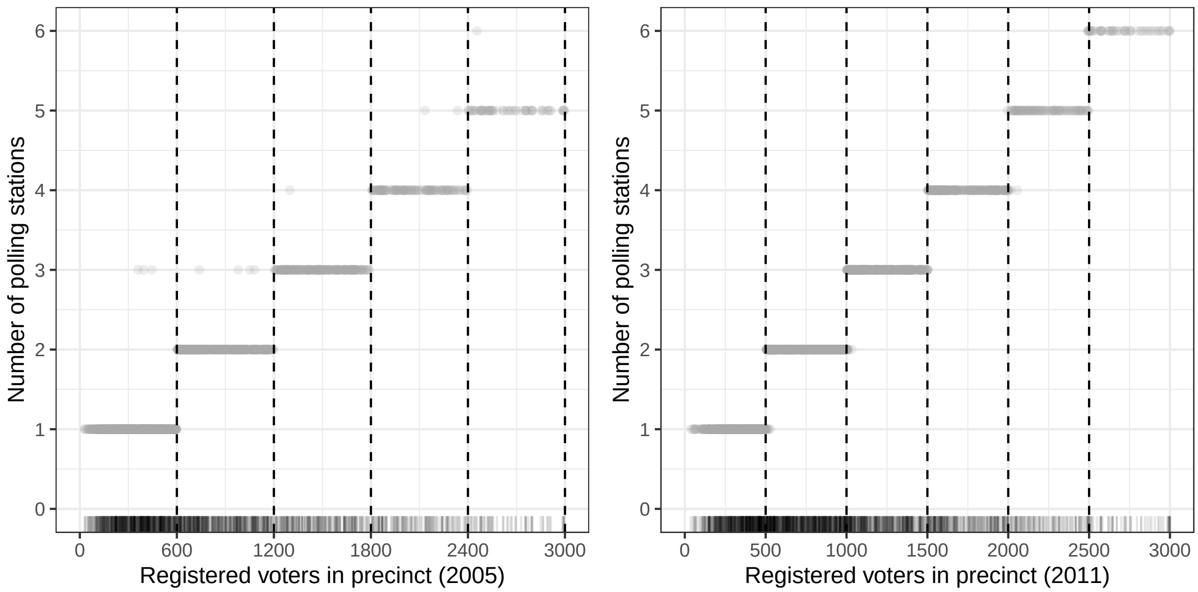


Figure 2: Number of polling stations vs. number of registered voters per precinct
Threshold for adding a new polling station at every 600 registered voters in a precinct in 2005 and every 500 registered voters in 2011

suggests near-perfect compliance with the threshold, but to account for imperfect compliance, we define the instrument for an additional station as follows:

$$\text{Predicted Split}_{iy} = \mathbb{1}\left(\frac{N_i}{T_y} > \text{nint}\left(\frac{N_i}{T_y}\right)\right),$$

where N_i is the number of registered voters in precinct i, T_y is the maximum number of registered voters per polling station in year y (500 or 600), and nint is a function returning the rounded closest integer. The treatment variable Split_{iy} is an indicator for whether precinct i does actually receive an additional polling station.

6.2.1 Validity of design

First, to justify the use of predicted split as an instrument for actual split, Table 1 presents first stage results. The strong coefficients on $\text{Predicted Split}_{iy}$, the predicted addition of a polling station, in combination with the near-zero control group means, suggests almost perfect compliance with the instrument.[4] In Section 7.5, we discuss and dismiss concerns about possible violations of the exclusion restriction.

Table 1: First Stage

	Split	
	2011	2005
	(1)	(2)
Predicted split	0.75***	1.00***
	(0.05)	(0.00)
Control Mean	0.09	0.00
Observations	188	81

Notes: *p<0.1; **p<0.05; ***p<0.01. Specifications are estimated with OLS within a 30-voter bandwidth with county fixed effects. Heteroskedasticity-robust standard errors in parentheses.

Second, the RDD framework also requires that there is no selective sorting of units into treatment and control groups. Figure 3 shows the density of units around the first four precinct-level thresholds in the two elections and visually suggests a lack of sorting. McCrary (2008) tests of the density of observations around the threshold fail to reject the null hypothesis of equal density on each side of the pooled threshold. This further supports the notion that identification is not confounded by the theoretical possibility of brokers' manipulation of the presence of an additional polling station.

Third, the RDD design requires that other covariates vary smoothly at the threshold. Appendix Tables A5 and A6 show balance tests for each election separately, where we run each variable as the outcome from our default regression specification and assess the significance of the treatment coefficient. Across both elections, there is very little evidence of imbalance in covariates. Consistent with chance, only 9 out of 105 coefficients are significantly imbalanced at the 10% level, with no consistent imbalance across elections. Importantly, precincts with an additional polling station in 2011 were not more pro-UP areas in 2005, which lessens concerns that the NEC could

[4] In our baseline specification, there is perfect compliance in 2005, as Table 1 suggests. As such, we do not need to instrument for Split_{iy} in the estimation for 2005.

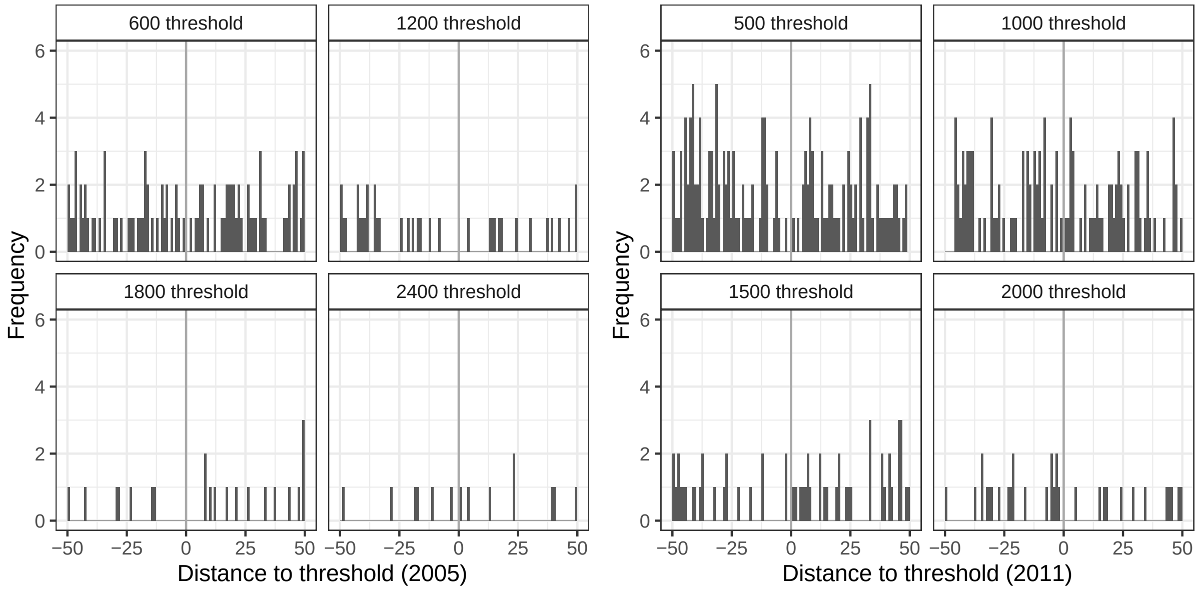


Figure 3: Density of precincts around thresholds where an additional polling station is added

have helped UP brokers manipulate the threshold. This, together with the evidence for lack of sorting, suggests an essentially random allocation of precincts into treatment and control across the elections.

6.2.2 Estimation

In our primary specifications, we restrict our sample to precincts that are within 30 voters of the threshold. Within this narrow bandwidth, as argued above, allocation into treatment and control is quasi-random. Under these conditions, we initially estimate the following equations in order to test for H1 and H2:

$$y_{ic} = \beta_1 \widehat{\text{Split}}_{ic} + \mu_c + \varepsilon_{ic}$$
$$\widehat{\text{Split}}_{ic} = \alpha_1 \text{Predicted Split}_i + \mu_c + \varepsilon_{ic},$$

where y_{ic} represents the relevant outcome variable in precinct i in county c, Predicted Split$_i$ indicates whether a given precinct is above or below the threshold, $\widehat{\text{Split}}_{ic}$ denotes the instrumented values of the treatment variable based on the first-stage regression, and μ_c are county-level fixed effects and account for differences in local campaigning due to concurrent Senate races. We use heteroskedasticity-robust standard errors. Separate estimations are used for each election. In Section 7.4, we show that the results are robust to (a) alternative choices of bandwidth, (b) the exclusion of county fixed effects, (c) the inclusion of linear trends on either side of the discontinuity.

Under our identification assumptions, β_1 represents the causal effect of an additional polling station. To test for H1, we use our various measures of polling station homogeneity as the outcome y_{ic}. To test for H2, we use electoral support for the UP and the CDC in 2005 and 2011 as outcomes. H3 suggests that there should be larger effects on electoral support for the UP in 2011 in precincts in which voters have to travel further to register and vote. To test that, we estimate the following specification:

$$y_{ic} = \beta_1 \widehat{\text{Split}}_{ic} + \beta_2 \text{Distance}_i + \beta_3 (\widehat{\text{Split}}_{ic} \times \text{Distance}_i) + \mu_c + \varepsilon_{ic}$$
$$\widehat{\text{Split}}_{ic} = \alpha_1 \text{Predicted Split}_i + \alpha_2 \text{Distance}_i + \alpha_3 (\text{Predicted Split}_i \times \text{Distance}_i) + \mu_c + \varepsilon_{ic},$$

where $Distance_i$ is the population-weighted distance of voters from their polling station in precinct i. We use a linear interaction for this distance measure. β_3 is the coefficient of interest.

7 Results

7.1 Polling station homogeneity

To provide evidence for H1, Table 2 shows the results of our baseline specification on our standardized measures of polling station homogeneity. Column 1, using a z-score index of the other outcomes, demonstrates that the polling station name similarity of voters significantly increases by 0.24 standard deviations (sd) when a precinct receives an additional polling station. Column 2 shows a 0.19 sd increase in exact last name matches, which is significant at the 10% level. Column 3 shows a 0.29 sd significant increase in the Soundex phonetic similarity measure. Column 4 suggests, albeit less precisely estimated, that the Levenshtein distance—where a lower distance represents more similar names—between registered voters decreases by 0.19 sd.

Table 2: Effects on Polling Station Homogeneity (H1)

		Name similarity measure		
	Index	Exact	Soundex	Lev.
	(1)	(2)	(3)	(4)
Split	0.24**	0.19*	0.29***	-0.19
	(0.11)	(0.10)	(0.11)	(0.12)
Observations	183	183	183	183

The outcomes are, in column (1), the standardized z-score index of name similarity measures; in column (2), the standardized average share of exact last name matches within polling stations in precinct; in column (3), the standardized average share of Soundex phonetic matches within polling stations in precinct; and in column (4), the standardized average Levenshtein distance between last names within polling stations in precinct.

All specifications are estimated within a 30-voter bandwidth. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Heteroskedasticity-robust standard errors in parentheses.

The results support H1 by consistently showing that a precinct being just above the threshold for receiving an additional polling station is associated with a significant and large increase in the last-name similarity of voters registered at each polling station within that precinct. This suggests that an additional polling station has meaningful effects on the homogeneity of voters within polling stations in a given precinct, which we consider to be an indicator of the likelihood that voters come from the same village under the influence of a particular broker. Given our argument about how the method by which brokers mobilize voters—by providing them with mass transportation—interacts with the sequential voter registration process, this evidence supports the argument that UP brokers mobilize voters *en masse* to signal their effort to higher-up brokers.

7.2 Party electoral support

To test for H2, Columns 1 and 2 of Table 3 provide estimates of our baseline equation for UP and CDC electoral support. Consistent with the hypothesis, the results suggest that an additional polling station is associated with a 6 percentage point increase in UP electoral support, which is both statistically significant and substantively large given the control group mean. Also consistent

Table 3: Effects on Party Electoral Support (H2)

	2011		2005	
	UP	CDC	UP	CDC
	(1)	(2)	(3)	(4)
Split	0.06**	0.03	-0.01	-0.00
	(0.03)	(0.02)	(0.02)	(0.02)
Control Mean	0.27	0.16	0.12	0.13
Observations	188	188	81	81

The outcomes are defined as precinct-level votes received by party divided by registered voters.

All specifications are estimated within a 30-voter bandwidth. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Heteroskedasticity-robust standard errors in parentheses.

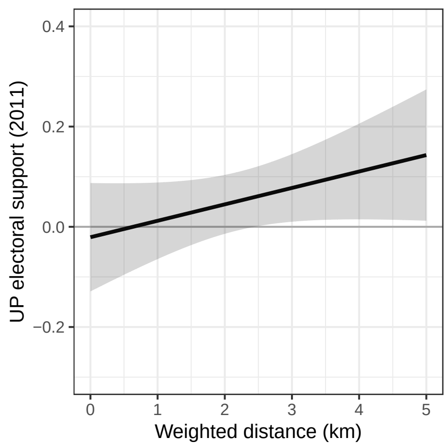
with H2 is the smaller and statistically insignificant effect on CDC electoral support.[5] Columns 3 and 4 of Table 3 provide the same set of results for the 2005 elections, which indicate that an additional polling station has no effect on electoral support for either party. The absence of any effects in 2005 supports H2 and aligns well with the qualitative evidence that neither party was able to mobilize voters in the first post-war election.

7.3 Party electoral support and distance to vote

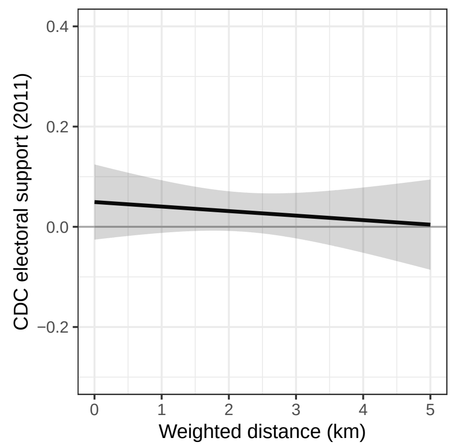
H3 implies that an additional station had a greater effect on UP electoral support in 2011 the further citizens were from their polling places. Figure 4 displays how the effect of an additional polling station varies as a function of the weighted average distance of precinct voters from their polling stations in the 2011 election.[6] Providing support for H3, the plots suggest that the significant treatment effect on UP electoral support is only observed for distances over approximately two kilometers, roughly the median distance observed in the data. This result is consistent with our qualitative evidence that the effect on UP electoral support in 2011 is driven by increased turnout buying rather than persuasion. Also consistent with both H2 and H3, there is no significant treatment effect across the domain of distance for CDC electoral support. Similar plots for

[5] In the Appendix, we show that this weak effect on CDC electoral support is fragile to bandwidth choice and, in Section 7.3, we demonstrate that it dissipates over distance.

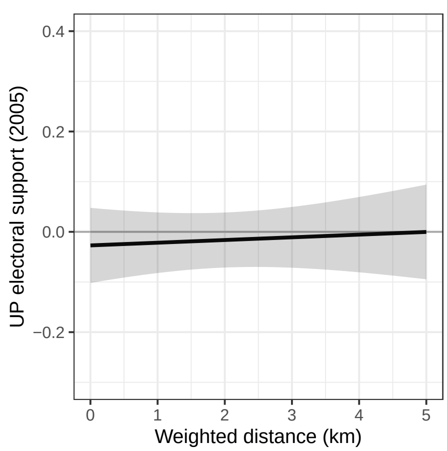
[6] The tables underlying these plots are shown in Appendix Table A7.



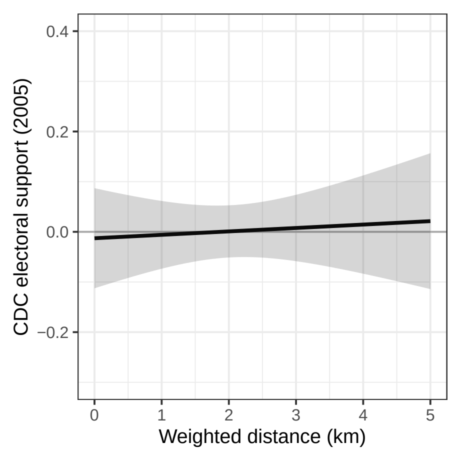
(a) UP electoral support (2011)



(b) CDC electoral support (2011)



(c) UP electoral support (2005)



(d) CDC electoral support (2005)

Figure 4: Marginal effect of an additional polling station on UP and CDC electoral support over distance
95% confidence intervals shaded in gray

2005 indicate no differential effect for either party across distance.

In summary, the electoral results suggest that an additional polling station capturing increased monitoring capacity is only associated with meaningful increases in UP electoral support in 2011 (H2), and that this electoral support is primarily captured through turnout buying (H3).

7.4 Robustness checks

In the Appendix, we present comprehensive robustness checks. Figure A1 provides visual evidence of the treatment effects. Figure A2 demonstrates the robustness of the results to varying the bandwidth, and Figure A3 to adding the forcing variable and its interaction with the treatment of an additional polling station. Table A8 modifies the specification of the homogeneity and electoral outcomes, by including all districts, including covariates as controls, removing county fixed effects, and conditioning voting outcomes on turnout. Figure A4 performs the same exercise for the distance interactions. Across the different permutations of robustness tests, we find consistent support for our results.

7.5 Alternative explanations

While we show evidence for the identifying assumption in Section 6.2.1, we rule out several alternative explanations here. First, it may be contended that, given the structural advantages accrued by the UP as an incumbent, the NEC registration process benefited UP strongholds. If this were the case, while we might expect overall higher electoral support for the UP, it would not affect the within-county estimates that we analyze. Further, there is no systematic evidence that the presence of an additional polling station was manipulated by affecting the number of voters who can register at particular precincts.

Second, the placement of precincts in 2011 may have been strategically closer to voters in UP strongholds. While theoretically feasible, the balance between whether a precinct is located in a UP area and the distance of citizens to their polling places in Table A5 provides evidence against this alternative explanation. Similarly, the balance on ethnic group variables across elections in Table A6 suggests that variation in the presence of these groups cannot account for results. Additionally, around two-thirds of the precincts in 2011 were within one kilometer of the nearest precinct in 2005, suggesting that their spatial allocation did not dramatically change with UP incumbency.

A third alternative explanation of our results relies on an argument about increased voting efficiency, which could affect electoral outcomes by either (a) changing vote choices at the ballot box or (b) increasing turnout by shortening voting lines. Neither can account, however, for: the absence of any effect in 2005, large effects for only the UP in 2011, and heterogeneous effects on the distance that voters need to travel to vote. First, the magnitude of the effects we observe is too large to be accounted for by a pure efficiency explanation given such high turnout rates—72% turnout in the first round of the 2011 election. Second, if voters change their mind at the ballot box, then an efficiency account would have to explain why nearly all of them switched to the UP and only did so in 2011. While unwarranted credit might be given to the incumbent for a smoother voting experience, either our baseline specification or the specification in which we condition on turnout suggest that an implausibly high proportion of voters would have to behave like this. Third, and most importantly, neither variant can account for why the effect varies with the distance voters must travel to vote. Individuals who have already traveled a long way are less likely to be affected by line length in deciding whether to vote. It is, therefore, implausible that increased voting efficiency affects these rural voters in ways that match the observed results.

8 Conclusion

The literature on broker monitoring has largely focused on the sophisticated machinery of institutionalized parties in well-established democracies. The dominant view is that such party structures do not exist in the young democracies of sub-Saharan Africa. In this paper, we contend that the reliance on sophisticated broker monitoring techniques is overstated. While young political parties may lack well-developed and institutionalized party structures, incumbent parties can exploit originally non-political networks and relationships to build a decentralized party structure to monitor and enforce broker performance on a large scale.

Specifically, weak parties can establish pyramidal structures, in which brokers at each level are responsible for using their own networks to monitor and incentivize brokers at the level below them. Through hierarchical decentralization, weakly institutionalized parties turn the challenge of monitoring brokers into multiple simple tasks. This strategy is primarily undertaken by incumbent parties because it hinges on access to state resources. Since brokers are initially drawn from outside the political party, they are unlikely to have strong ideological affinity with it and are instead incentivized by the potential rewards for good performance. Thus, a political party can only be successful in using such a pyramidal structure for large-scale voter mobilization when it has access to the resources required to monitor and incentivize broker effort.

We study the 2005 and 2011 elections in post-war Liberia, a hard case with barely a decade of democratic experience and a weak party system. We show that the then-incumbent Unity Party relied on such a pyramidal structure to monitor the efforts of low-level brokers during voter registration and election periods. Rich qualitative evidence suggests that the resources afforded by incumbency enabled the UP to develop such a structure. Moreover, we use administrative data and particular features of Liberian electoral administration to provide causal evidence that low-level brokers mobilize voters *en masse* to signal effort to higher-level brokers. We further show that increased monitoring ability of higher-level brokers improves the incumbent party's electoral performance, particularly in precincts in which there is greater scope for turnout buying.

Our argument and the evidence we provide have important implications not only for the literature on voter mobilization in sub-Saharan Africa, but also for democratic consolidation in less-developed democracies. As long as weakly institutionalized parties are able to develop decentralized party structures for large-scale voter mobilization through brokers, these democracies may face difficulties with party institutionalization and consolidation. One implication of our results is that reducing the opportunities for brokers to signal effort may render monitoring harder to sustain. Consequently, electoral reforms, such as permanent voter registration systems, may undermine the ability of parties to monitor their brokers. Future research on the impact of such reforms is warranted.

References

Adejumobi, Said. 2006. "Reviving a Failed State: the 2005 General Elections in Liberia." *Journal of African Elections* 5(1):126–151.

Afrobarometer. 2012. "Afrobarometer Data, Liberia, Round 5, 2012.".
URL: *http://afrobarometer.org/data/320*

Aspinall, Edward. 2014. "When Brokers Betray: Clientelism, Social Networks, and Electoral Politics in Indonesia." *Critical Asian Studies* 46(4):545–570.

Baldwin, Kate. 2013. "Why Vote with the Chief? Political Connections and Public Goods Provision in Zambia." *American Journal of Political Science* 57(4):794–809.

Bates, Robert H. 1974. "Ethnic Competition and Modernization in Contemporary Africa." *Comparative Political Studies* 6(4):457.

Blattman, Chris, Horacio Larreguy, Benjamin Marx and Otis Reid. 2019. "Eat Widely, Vote Wisely? Lessons from a Campaign Against Vote Buying in Uganda." *Working Paper. https://www.dropbox.com/s/nlcn1gomlls2yxo/Manuscript_Final.pdf?dl=0* .

Boas, Morten and Mats Utas. 2014. "The Political Landscape of Postwar Liberia: Reflections on National Reconciliation and Elections." *Africa Today* 60(4):47–65.

Bratton, Michael. 2008. "Vote Buying and Violence in Nigerian Election Campaigns." *Electoral Studies* 27(4):621–632.

Brierley, Sarah and Noah Nathan. 2019. "The Connections of Party Brokers." *Working paper. http://pelg.ucsd.edu/Brierley_2019.pdf* .

Calvo, Ernesto and Maria Victoria Murillo. 2004. "Who Delivers? Partisan Clients in the Argentine Electoral Market." *American Journal of Political Science* 48(4):742–757.

Carter Center. 2011. "National Elections in Liberia." *Carter Center* .
URL: *https://www.cartercenter.org/resources/pdfs/news/peace_publications/election_reports/liberia2011-finalrpt.pdf*

Carter Center and National Democratic Institute. 2005. "Liberia Runoff Election: Background Information." *Carter Center* 4.

URL: *https://www.cartercenter.org/documents/2204.pdf*

Casey, Katherine. 2015. "Crossing Party Lines: The Effects of Information on Redistributive Politics." *The American Economic Review* 105(8):2410–2448.

Conroy-Krutz, Jeffrey. 2013. "Information and Ethnic Politics in Africa." *British Journal of Political Science* 43(2):345–373.

De Kadt, Daniel and Horacio Larreguy. 2018. "Agents of the Regime? Traditional leaders and Electoral Behavior in South Africa." *Journal of Politics* 80(2):382–399.

Duarte, Raul, Frederico Finan, Horacio Larreguy and Laura Schechter. 2019. "Networks, Information and Vote Buying." *Working paper. https://www.dropbox.com/s/h06k32c9mlmc1yf/Paraguay_v41.pdf?dl=0 .*

Finan, Frederico and Laura Schechter. 2012. "Vote-Buying and Reciprocity." *Econometrica* 80(2):863–881.

Gingerich, Daniel W. and Luis Fernando Medina. 2013. "The Endurance and Eclipse of the Controlled Vote: A Formal Model of Vote Brokerage Under the Secret Ballot." *Economics and Politics* 25(3):453–480.

Gottlieb, Jessica and Horacio Larreguy. 2019. "An Informational Theory of Electoral Targeting in Young Clientelistic Democracies: Evidence from Senegal." *Quarterly Journal of Political Science, forthcoming* .

Harris, David. 2006. "Liberia 2005: An Unusual African Post-conflict Election." *The Journal of Modern African Studies* 44(03):375–395.

Ichino, Nahomi and Noah L Nathan. 2013. "Crossing the Line: Local Ethnic Geography and Voting in Ghana." *American Political Science Review* 107(02):344–361.

Kitschelt, Herbert and Daniel M Kselman. 2013. "Economic Development, Democratic Experience, and Political Parties' Linkage Strategies." *Comparative Political Studies* 46(11):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. Herbert Kitschelt and Steven I. Wilkinson. Cambridge University Press pp. 1–47.

Koter, Dominika. 2013. "King Makers: Local Leaders and Ethnic Politics in Africa." *World Politics* 65(2):187–232.

Kramon, Eric. 2016. "Electoral Handouts as Information: Explaining Unmonitored Vote Buying." *World Politics* 68(3):454–498.

Larreguy, Horacio. 2013. "Monitoring Political Brokers: Evidence from Clientelistic Networks in Mexico.".
URL: `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2225027`

Larreguy, Horacio, Cesar E. Montiel Olea and Pablo Querubin. 2017. "Political Brokers: Partisans or Agents? Evidence from the Mexican Teacher's Union." *American Journal of Political Science* 61(4):877–891.

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." *American Political Science Review* 100(1):160–179.

Magaloni, Beatriz. 2006. *Voting for Autocracy*. New York: Cambridge University Press.

Mainwaring, Scott. 1999. *Rethinking party systems in the third wave of democratization: the case of Brazil*. Stanford University Press.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142(2):698–714.

Nathan, Noah L. forthcoming. "Electoral Consequences of Colonial Invention: Brokers, Chiefs, and Distribution in Northern Ghana." *World Politics* .

Nichter, Simeon. 2008. "Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot." *American Political Science Review* 102(1):19–31.

Novaes, Lucas M. 2015. Modular Parties: Party Systems with Detachable Clienteles. PhD thesis University of California, Berkeley. `http://digitalassets.lib.berkeley.edu/etd/ucb/text/Novaes_berkeley_0028E_15846.pdf`.

Posner, Daniel N. 2005. *Institutions and Ethnic Politics in Africa*. Cambridge University Press.

Stokes, Susan C. 2005. "Perverse Accountability: A Formal Model of Machine Politics with Evidence from Argentina." *American Political Science Review* 99(3):315–325.

Stokes, Susan C., Thad Dunning, Marcelo Nazareno and Valeria Brusco. 2013. *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics*. Cambridge University Press.

Szwarcberg, Mariela. 2014. "Political parties and rallies in Latin America." *Party Politics* 20(3):456–466.

Thachil, Tariq. 2011. "Embedded Mobilization: Nonstate Service Provision as Electoral Strategy in India." *World Politics* 63(3):434–69.

Thachil, Tariq. 2014. "Elite Parties and Poor Voters: Theory and Evidence from India." *American Political Science Review* 108(2):454–477.

USAID. 2015. "Liberia Electoral Access and Participation (LEAP) Survey." *USAID* .
URL: *http://pdf.usaid.gov/pdf_docs/PA00KJW1.pdf*

Vicente, Pedro C. 2014. "Is Vote Buying Effective? Evidence from a Field Experiment in West Africa." *Economic Journal* 124:F356–F387.

Vicente, Pedro C. and Leonard Wantchekon. 2009. "Clientelism and Vote Buying: Lessons from Field Experiments in African Elections." *Oxford Review of Economic Policy* 25(2):292–305.

Zarazaga, Rodrigo. 2014. "Brokers beyond clientelism: A new perspective through the Argentine case." *Latin American Politics and Society* 56(3):23–45.

A Online appendix for 'How Weakly Institutionalized Parties Monitor Brokers in Developing Democracies: Evidence from Post-conflict Liberia"

Contents

A.1 Qualitative data collection

We collected qualitative evidence for this paper, through interviews and focus groups, from June-August 2016 and July-August 2017 as part of broader projects on the civil war and on post-war political development in Liberia. In 2016, our fieldwork sites were in Montserrado and Margibi counties. In 2017, we added further evidence from Montserrado county and new evidence from Bomi County.

A.1.1 Interviews

Our primary interviewees comprised officials from the National Elections Commission (NEC), the Unity Party (UP), and the Congress for Democratic Change (CDC). Interviewees were targeted based on their seniority within their respective organizations and their ability to speak about election campaigning, the voter registration period, and election-day procedures. In 2016, we interviewed relatively senior members from the NEC, UP, and CDC party structure, on campaign strategy, and their use of electoral results to monitor intermediate- and low-tier brokers. In 2017, we added a set of lower-level representatives from UP and CDC involved in low-tier broker monitoring, and representatives from several NGOs and international organizations that work with the political parties.

A.1.2 Focus groups

We conducted six civilian focus groups in 2016 to discuss political party practices during campaigning period, voter registration, and on election day. Focus groups consisted of between six to ten individuals, and were conducted in six different communities across Monrovia based on 1) broader project constraints[7] and 2) the local community leader's willingness to convene participants for our study. The six communities are geographically dispersed across the city. To recruit focus group participants, we first identified our targeted communities and spoke to the local community leader about conducting a focus group in the community. Having obtained their permission, the local community leader then assisted us with identifying individuals who qualified for our study.

7 Chiefly, focus groups were targeted at individuals who were at least seven years old at the start of the first civil war, and were not an ex-combatant who participated in the civil war.

A.2 Name similarity measures and validation tests

This section explains and validates the last name similarity measures used as a measure of polling station homogeneity in the paper, as described in Section 6.1.1. Liberia has strong traditions of patrilocality which render last name similarity a potentially useful measure of villages of origin. The key assumption, as highlighted there, is that measuring the similarity of last names registered at a given polling station within a precinct is a reasonable measure of the likelihood that registered individuals come from the same village.

A.2.1 Measures

We employ three measures of last name similarity: exact matches, Soundex, and Levenshtein distance. These measures each have different, related properties. Exact matches, intuitively, measure the incidence of exact last name spelling matches in the voter register. If last name i exactly matches last name j this takes a value of 1 and 0 otherwise. However, in contexts such as Liberia, where relatively few citizens possess formal identity documents to reference name spelling and literacy remains low, relying on exact matching is likely to underestimate the incidence of related individuals registered at a given polling station.

Soundex measures, instead, measure the phonetic similarity of matches by encoding words into phonetic codes and assessing binary matches of this phonetic code. Last names which are spelled differently but pronounced relatively similarly, therefore, will register as a Soundex phonetic match. If i is a phonetic match for j then we code 1 for the Soundex match, and 0 otherwise.

The Levenshtein distance accounts, instead, for minor variations in spelling. Specifically, it measures the number of individual character deletions, insertions, or substitutions that are required to transform one last name into another. We normalize the Levenshtein distance by the respective number of characters in last name i. If i requires no edits to turn into j then it takes a Levenshtein distance of 0, and higher values indicate more dissimilar names.

For an illustration of the measures, suppose a polling station contains three voters with surnames recorded as Kollie, Kolleh, and Konneh. Table A1 presents the matrix of name similarity measures where the three entries in cell (i, j) indicate whether i is an exact match for j, whether i is a Soundex match for j, and the normalized Levenshtein distance between i and j. None of the names are exact matches, while (Kollie, Kolleh) is a phonetic match, and (Kollie, Kolleh) has the same Levenshtein distance as (Kolleh, Konneh) but less than (Kollie, Konneh). In this simple example, averaging across each measure for each individual, the polling station measures of last name similarity would be (0, 0.33, 0.44). In the analysis, we log all these aggregated variables, to account for skewness, and standardize them to have mean 0 and standard deviation 1.

Table A1: Illustration of Name Similarity Measures

	Kollie	Kolleh	Konneh	Exact	Soundex	Lev.
Kollie	-	(0,1,0.33)	(0,0,0.67)	0	0.5	0.5
Kolleh	(0,1,0.33)	-	(0,0,0.33)	0	0.5	0.33
Konneh	(0,0,0.67)	(0,0,0.33)	-	0	0	0.5
				0	**0.33**	**0.44**

A.2.2 Validation

We validate the use of last name similarity as a proxy for the homogeneity of polling stations, in terms of the locations from which registered voters come, in two ways. First, we draw on census data using the approach outlined in Appendix A.3 to estimate the level of ethnic homogeneity at the voting precinct level. We proxy for ethnic homogeneity using a Hirschman-Herfindahl index (HHI), where higher values represent *less* ethnic diversity at the precinct level. In Panel A of Table A2, we regress each of the name similarity measures onto this ethnic homogeneity measure. Each of the measures, and the index, correlates very strongly with ethnic homogeneity: a one standard deviation (sd) increase in the precinct's imputed ethnic homogeneity is associated with a 0.38 sd increase in the last name similarity index of the voters registered at the polling stations within that precinct. This provides validation that last name similarity represents a measure of homogeneity.

Table A2: Validating Name Similarity Measures

		Name similarity measure		
	Index (1)	Exact (2)	Soundex (3)	Lev. (4)
A. Correlation with Ethnicity HHI				
Ethnicity HHI	0.38***	0.42***	0.42***	-0.21***
	(0.03)	(0.03)	(0.03)	(0.03)
Observations	1410	1410	1410	1410
B. HIES Name Similarity Validation				
Within-EA	0.23***	0.25***	0.20***	-0.14***
	(0.05)	(0.05)	(0.05)	(0.05)
Observations	856	856	856	856

Notes: Dependent variables are standardized. Panel A: Ethnicity HHI measure is standardized. Specification estimated with county fixed effects. Panel B: Within-EA is an indicator where the across-EA measure is the excluded category. Specification estimated with EA fixed effects. *p<0.1; **p<0.05; ***p<0.01. Heteroskedasticity-robust standard errors in parentheses.

Second, we provide direct evidence that last names are more similar *within* localities than *across* them by drawing on the 2016 Household and Expenditure Survey (HIES), a nationally representative consumption survey of over 8,000 households. Using a restricted-access version of this dataset generously shared with us by the statistical authority, we extract the last name of each household head and the geolocation of the enumeration area (EA). Each EA, mapping relatively closely to a specific locality, contains 10 households. We then compute versions of the last name similarity measures at the EA level as follows:

1. Within EA i we compute the last name similarity measures between all household heads in i and aggregate to the EA level.
2. We geographically match each EA i to the closest other EA j and compute the last name similarity measures between all household heads in i and those in j, and aggregate to the EA level.

To map onto the empirical setting of how far citizens must travel to vote, we discard all EAs more than 5 km from another EA. This restriction leaves 428 EAs at a mean distance between EAs of 1.92 km, which is roughly the mean distance that citizens need to commute in our sample to cast their vote. We then regress the last name similarity measures onto an indicator for whether the similarity measure is *within EA* as opposed to *across EA*, as well as EA fixed effects so that to only compare similarity measures within each EA. Panel B of Table A2 presents results from this demanding test. Household head last name similarity measures are 0.23 sd greater when comparing last names *within* EA to the last names of household heads in the closest EA. This supports that names are more similar within, versus across, nearby localities.

A.3 Precinct-level variables

Here, we provide more detail on how the weighted precinct-level measures were constructed for both the key distance variable and other population variables. We use geolocated 2008 census data from the Liberia Institute of Statistics and Geo-Information Services (LISGIS), combined with the coordinates of all precincts in the two elections from the United Nations Mission in Liberia (UNMIL), to calculate the distance from each census population location to the closest precinct in the same district.

Having matched each population location to the closest precinct, we assign the population of eligible voters to that precinct and then, at the precinct level, take a weighted average of the distance each assigned population location is from the precinct. In general, this parsimonious approach performs well, however especially in more urban areas there are relatively few distinct population locations in the census for a substantial number of precincts. This means that many of the precincts aren't closest to any population location and so have no voting populations assigned to them. In such cases where there is no population assigned to the precinct, we impute the measure by assigning the distance measure from the closest precinct that does have any population assigned to it. Given that precincts with no assigned populations tend to be closer than those with assigned populations, we believe this is a reasonable method of imputing the distance measure where necessary. For the other socio-economic variables on which we check balance, we follow the same basic approach and also weight on population.

Then, to assess how reasonable this approach was, we tabulated the share of precincts in each district with data imputed in this way. Table A3 shows the proportion of precincts across particular districts where we had to impute the distance measure. This table presents 2011 data, but 2005 is very similar. Only the 28 (out of 136) districts with any imputed data are listed, and the remaining 108 districts have no imputed data at all. As is clear, Greater Monrovia and Commonwealth-B (representing Buchanan) have by far the most imputed values. Thus, for the bulk of the analysis, we exclude just these districts. If we used the full sample, 475/1780 (26.7%) of the precincts in 2011 would be using imputed data for the distance variable. Excluding these two districts means that 107/1409 (7.6%) of the precincts are using imputed measures for the distance variable. Within our baseline bandwidth of 30 registered voters, including these districts means that 69/236 (29.2%) observations in 2011 are using imputed distance values while excluding them means that 11/178 (6.1%) observations are.

Table A3: Share of Imputed Data for Top Districts

District	Number of precincts	Share imputed
Greater Monrovia	355	1.00
Commonwealth-B	15	0.93
St. Paul River	39	0.51
Gee	2	0.50
Garr-Bain	23	0.39
Firestone	34	0.38
Yarmein	12	0.33
Careysburg	18	0.28
Harper	22	0.27
Kakata	54	0.26
Mambah Kaba	28	0.21
Senjeh	25	0.20
Tchien	15	0.20
Jorquelleh	32	0.19
Greenville	6	0.17
Trenbo	6	0.17
Wee-Gbehyi-Mahn	12	0.17
Vahun	7	0.14
Meinpea-Mahn	12	0.08
Doe	13	0.08
Sanniquellie Mahn	13	0.08
Gibi	14	0.07
Suakoko	15	0.07
Yeallequelleh	18	0.06
Pleebo/Sodoken	20	0.05
Neekreen	21	0.05
Zorzor	26	0.04
Voinjama	28	0.04

A.4 Descriptive Statistics

Table A4: Descriptive Statistics

	Mean	SD	Min	Max	N
A. Polling station homogeneity					
Number of polling stations	2.36	1.31	1.00	9.00	1410
Registered voters	930.88	633.47	45.00	3995.00	1410
Exact match	0.02	0.02	0.00	0.20	1410
Soundex match	0.03	0.02	0.01	0.21	1410
Levenshtein distance	0.98	0.04	0.72	1.09	1410
B. 2011 election					
Number of polling stations	2.22	1.24	1.00	7.00	1409
Registered voters	871.88	592.90	38.00	3364.00	1409
UP electoral support	0.29	0.16	0.00	0.84	1409
CDC electoral support	0.18	0.14	0.00	0.74	1409
Distance to precinct (km)	2.03	1.41	0.00	14.01	1409
C. 2005 election					
Number of polling stations	1.79	1.02	1.00	8.00	1153
Registered voters	772.78	594.68	30.00	4523.00	1153
UP electoral support	0.10	0.13	0.00	0.71	1153
CDC electoral support	0.15	0.17	0.00	0.88	1153
Distance to precinct (km)	2.78	2.39	0.02	33.93	1153

Notes: For reasons outlined in Appendix Section A.3 we exclude Monrovia and Buchanan districts in all the analysis.

A.5 Covariate balance

Table A5: Covariate Balance

	2011		2005	
	Coef. (1)	SE (2)	Coef. (3)	SE (4)
Weighted distance	-0.10	(0.16)	-0.22	(0.15)
Share urban	-0.06	(0.13)	0.16	(0.13)
Share with furniture	-0.02	(0.12)	0.01	(0.15)
Share with mattress	-0.03	(0.15)	0.16	(0.19)
Share with radio	0.18	(0.15)	0.08	(0.19)
Share with television	0.13	(0.10)	0.06	(0.08)
Share with cell phone	0.11	(0.14)	0.06	(0.18)
Share with motorcycle	-0.16	(0.29)	0.08	(0.12)
Share with vehicle	-0.02	(0.11)	0.15	(0.09)
Share with fridge	0.06	(0.11)	0.12	(0.09)
Share grows rice	0.12	(0.12)	0.14	(0.16)
Share grows cassava	-0.04	(0.13)	0.25	(0.20)
Share grows plantain	-0.21	(0.17)	0.33*	(0.19)
Share grows rubber	-0.03	(0.20)	0.14	(0.27)
Share grows palm oil	-0.20	(0.21)	0.06	(0.19)
Share grows coffee	-0.14	(0.13)	0.21*	(0.11)
Share grows cocoa	-0.24*	(0.14)	0.15	(0.10)
Share grows coconut	-0.48**	(0.25)	0.40	(0.25)
Share has livestock	-0.21	(0.25)	0.13	(0.17)
Share has poultry	0.19	(0.19)	0.60**	(0.30)
Share has fishery	-0.21	(0.22)	-0.04	(0.10)
Share has other ag.	0.08	(0.19)	0.02	(0.12)
Share male	-0.18	(0.15)	0.45	(0.47)
Mean age	0.14	(0.16)	-0.02	(0.22)
Share Christian	0.16	(0.13)	-0.01	(0.17)
Share Liberian	-0.04	(0.15)	0.44	(0.45)
Mean residence length	0.12	(0.12)	-0.16	(0.15)
Share displaced	0.10	(0.15)	0.14	(0.19)
Share resettled	0.02	(0.14)	0.12	(0.18)
Share literate	-0.05	(0.14)	0.23	(0.20)
Share without education	0.07	(0.17)	-0.17	(0.18)
Share student	0.07	(0.17)	0.07	(0.16)
Share has occupation	-0.04	(0.18)	-0.11	(0.24)
UP electoral support (2005)	-0.14	(0.16)		

All outcomes are at the precinct level. The outcomes are standardized measures drawn from the 2008 National Population and Housing Census.
All specifications are estimated within a 30-voter bandwidth. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Heteroskedasticity-robust standard errors in parentheses.

Table A6: Covariate Balance (Ethnic groups)

	2011		2005	
	Coef. (1)	SE (2)	Coef. (3)	SE (4)
Ethnic homogeneity	0.08	(0.11)	0.04	(0.15)
Share Kpelle	-0.02	(0.08)	0.05	(0.18)
Share Bassa	-0.10	(0.10)	0.02	(0.16)
Share Grebo	0.02	(0.12)	-0.03*	(0.02)
Share Gio	-0.19	(0.17)	0.12	(0.23)
Share Mano	0.15	(0.16)	-0.16	(0.23)
Share Kru	-0.18	(0.20)	-0.01	(0.03)
Share Loma	-0.03	(0.17)	-0.11	(0.14)
Share Kissi	0.28**	(0.14)	0.12	(0.11)
Share Gola	-0.13	(0.12)	0.22	(0.20)
Share Vai	-0.05	(0.07)	-0.27	(0.19)
Share Krahn	0.01	(0.01)	0.01	(0.03)
Share Mandingo	-0.26	(0.26)	0.06	(0.04)
Share Gbandi	0.03	(0.21)	0.04*	(0.02)
Share Mende	0.19	(0.21)	-0.07	(0.23)
Share Sapo	0.29*	(0.17)	-0.05	(0.05)
Share Belle	0.00	(0.01)	0.06	(0.04)
Share Dey	-0.03	(0.04)	-0.28	(0.28)
Share other Liberian	-0.13	(0.12)	0.00	(0.02)

All outcomes are at the precinct level. The outcomes are standardized measures drawn from the 2008 National Population and Housing Census. Ethnic groups listed from largest to smallest.
All specifications are estimated within a 30-voter bandwidth. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Heteroskedasticity-robust standard errors in parentheses.

A.6 Distance interactions

Table A7: Effects on Party Electoral Support Conditional on Distance (H3)

	2011		2005	
	UP	CDC	UP	CDC
	(1)	(2)	(3)	(4)
Split	0.04	0.02	-0.02	-0.00
	(0.03)	(0.02)	(0.02)	(0.02)
Split × Distance	0.07**	0.01	0.02	-0.01
	(0.03)	(0.02)	(0.03)	(0.03)
Control Mean	0.27	0.16	0.12	0.13
Observations	188	188	81	81

The outcomes are party electoral support in the different elections. *Distance* is the standardized weighted distance of voters to a given precinct.

All specifications are estimated within a 30-voter bandwidth. * p < 0.1, ** p < 0.05, *** p < 0.01. Heteroskedasticity-robust standard errors in parentheses.

A.7 Regression discontinuity plots

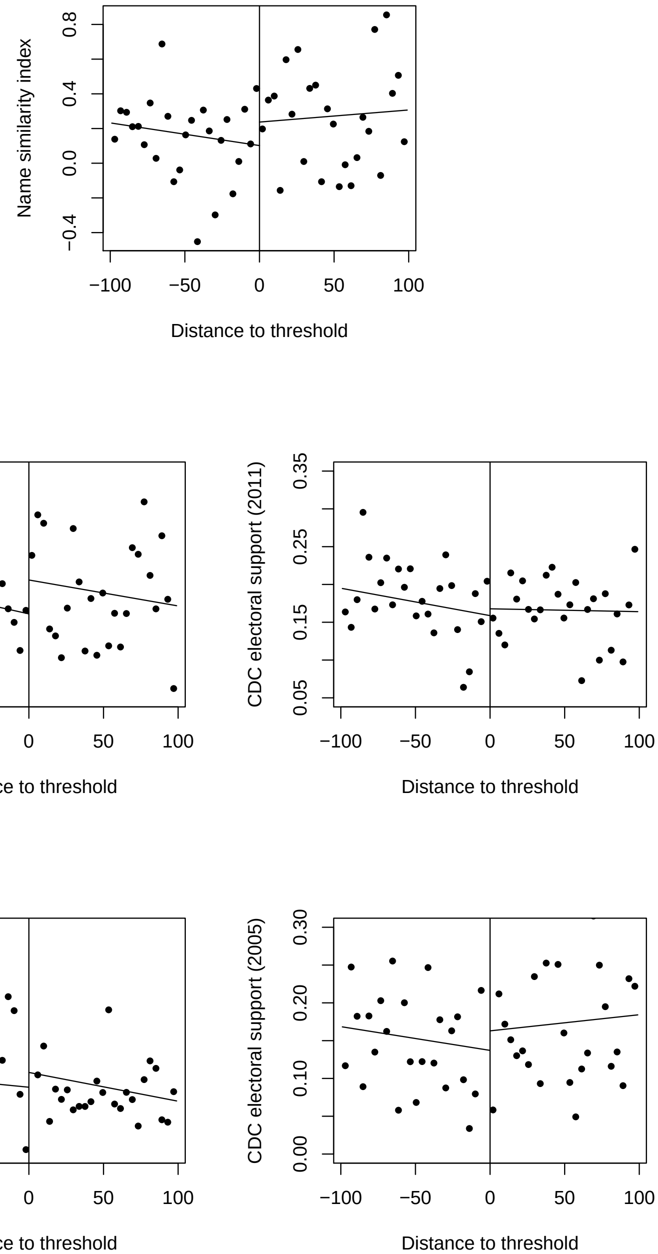
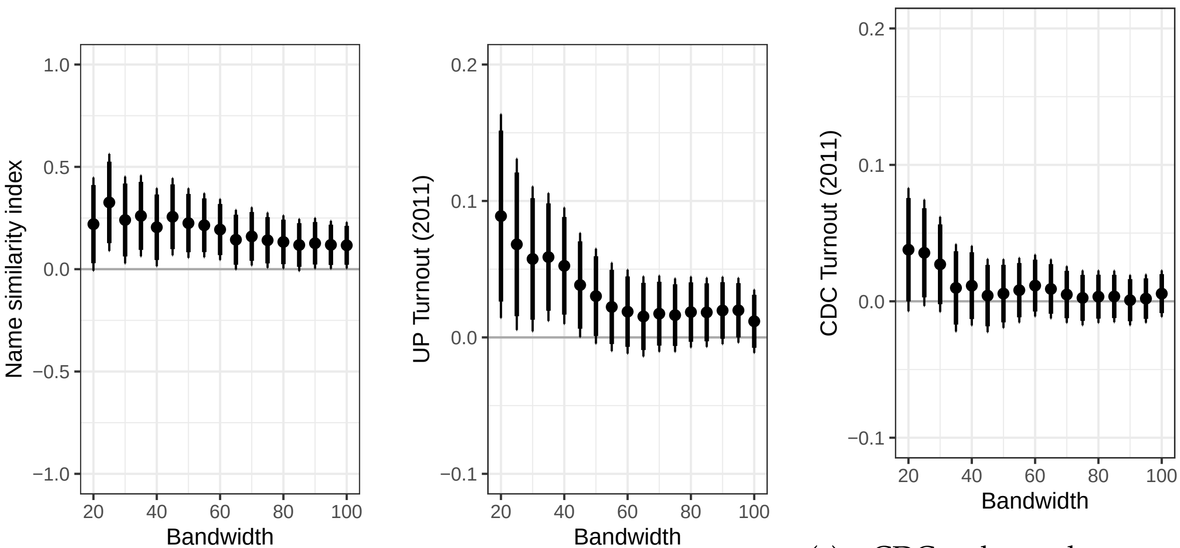


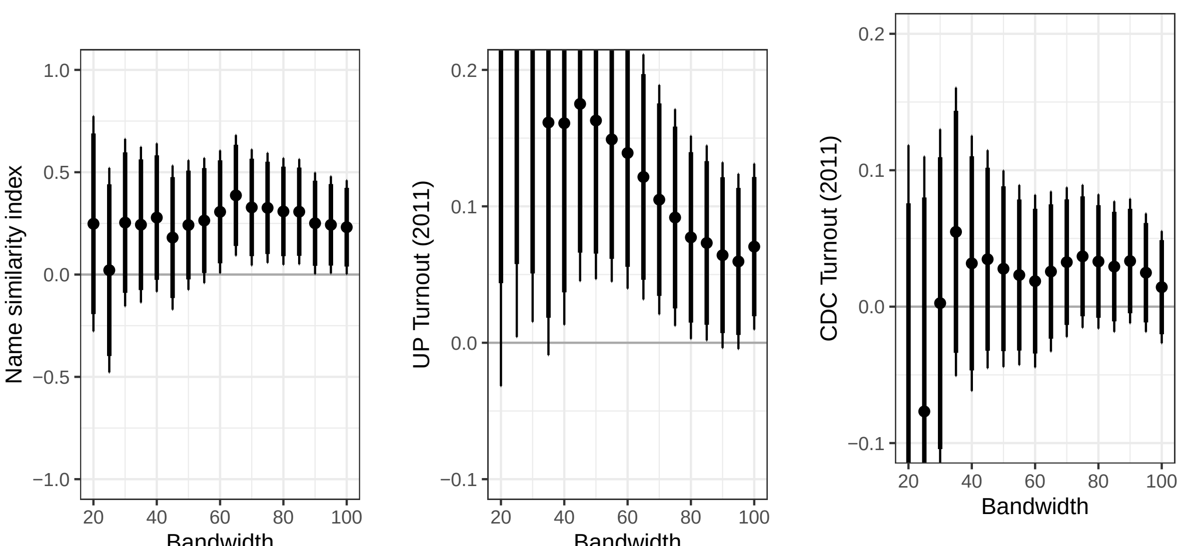
Figure A1: Change in outcomes around the threshold where an additional polling station is added

A.8 Robustness tests



(a) Name similarity index (b) UP electoral support (2011) (c) CDC electoral support (2011)

Figure A2: Varying bandwidths for the marginal effect of an additional polling station on outcomes (Baseline specification)
90% and 95% confidence intervals plotted for each bandwidth



(a) Name similarity index (b) UP electoral support (2011) (c) CDC electoral support (2011)

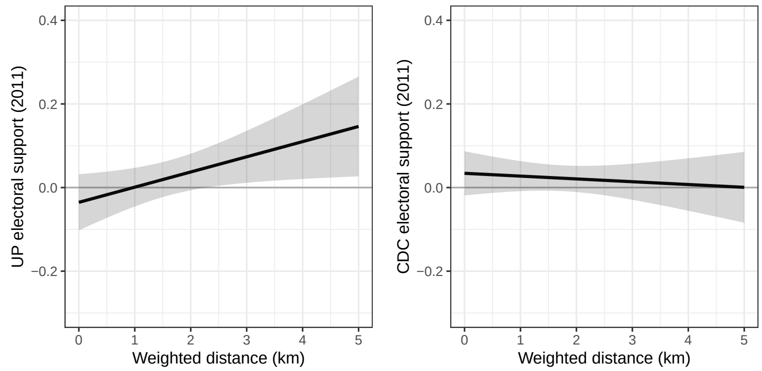
Figure A3: Varying bandwidths for the marginal effect of an additional polling station on outcomes (Interacting treatment with linear proximity to threshold)
90% and 95% confidence intervals plotted for each bandwidth

Table A8: Robustness Tests (H1 & H2)

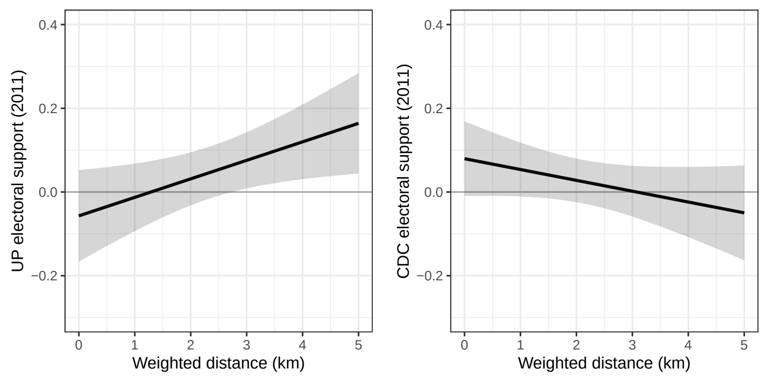
	Homogeneity outcomes				Electoral outcomes			
		Name similarity measure			2011		2005	
	Index	Exact	Soundex	Lev.	UP	CDC	UP	CDC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. All districts								
Split	0.23**	0.18*	0.28***	-0.18*	0.04*	0.02	-0.01	0.00
	(0.10)	(0.10)	(0.10)	(0.11)	(0.02)	(0.01)	(0.02)	(0.02)
Control Mean	-0.14	-0.12	-0.18	0.10	0.28	0.20	0.14	0.16
Observations	228	228	228	228	248	248	101	101
B. Including covariates								
Split	0.23**	0.18*	0.27***	-0.18*	0.06**	0.03	-0.02	-0.02
	(0.10)	(0.10)	(0.10)	(0.11)	(0.03)	(0.02)	(0.02)	(0.02)
Control Mean	0.12	0.17	0.10	-0.05	0.27	0.16	0.12	0.13
Observations	183	183	183	183	188	188	81	81
C. No county FE								
Split	0.24*	0.22*	0.32**	-0.14	0.04	0.01	-0.02	0.04
	(0.14)	(0.13)	(0.13)	(0.16)	(0.03)	(0.03)	(0.03)	(0.04)
Control Mean	0.12	0.17	0.10	-0.05	0.27	0.16	0.12	0.13
Observations	183	183	183	183	188	188	81	81
D. Electoral support conditioned on turnout								
Split					0.08**	0.02	-0.02	-0.01
					(0.03)	(0.02)	(0.03)	(0.03)
Control Mean					0.39	0.23	0.17	0.19
Observations					188	188	81	81

Panel A includes all districts; Panel B controls for standardized first PCA indices of covariates relating to assets, agriculture, demography, literacy and ethnic composition; Panel C removes county fixed effects from estimating equation; Panel D considers as an outcome party electoral support over total votes rather than registered voters.

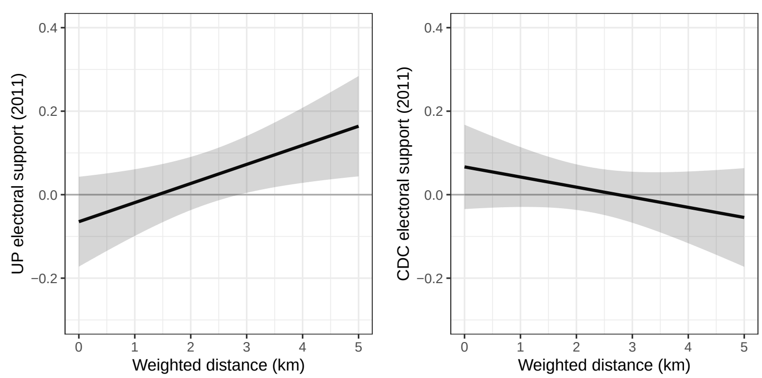
All specifications are estimated within a 30-voter bandwidth. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Heteroskedasticity-robust standard errors in parentheses.



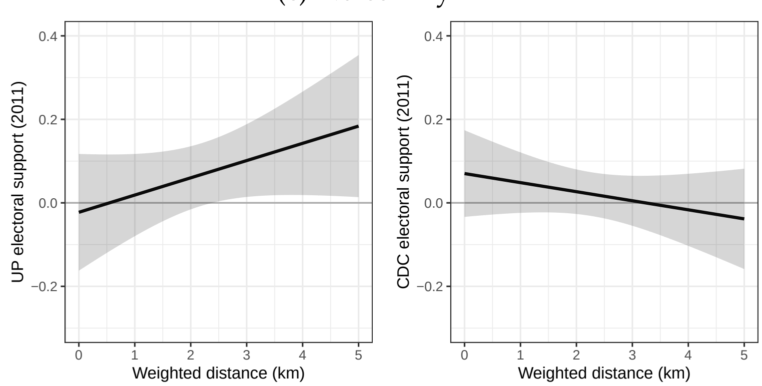
(a) All districts



(b) Including covariates



(c) No county FE



(d) Electoral support conditioned on total votes

Figure A4: Robustness of marginal effect of an additional polling station on UP and CDC electoral support over distance (H3)