Ethnicity and Electoral Fraud in New Democracies: Modelling Political Party Agents in Ghana

Nahomi Ichino *   Jake Bowers †   Mark M. Fredrickson ‡

October 5, 2013 (Version: f300d33)

Abstract

How might ethnicity affect political party strategies for electoral fraud and complicate the consolidation of new democracies? We engage this question by modelling the efforts of political party agents to inflate the voters register ahead of the 2008 general elections in Ghana. Because we cannot directly observe rigging of the voter registration process, we take advantage of a randomized field experiment that placed observers at registration centers to affect the behavior of the party agents. We propose two simple models of party agent behavior, one of which incorporates an additional preference for visiting registration centers in ethnically homogeneous areas. We then use new network data to assess the evidence against these models given the design and data from the field experiment. In contrast to most common analyses of randomized experiments, our aim in this paper is not to learn about a scalar effect of treatment, but to learn about models of treatment propagation.

*Associate Professor, Department of Government, Harvard University (`nichino@gov.harvard.edu`). *Acknowledgements:* Thanks to the participants in the Causality in Political Networks conference, the workshop on methodology in the Department of Government at the University of Texas, Jeff Conroy-Krutz at APSA 2013, and the SAMSI Workshop on Social Networks 2013 for helpful advice and the Ghana Statistical Service and the Republic of Ghana Ministry of Roads and Highways for data.

†Associate Professor, Department of Political Science, University of Illinois @ Urbana-Champaign (`jwbowers@illinois.edu`)

‡Ph.D. Candidate, Department of Political Science, University of Illinois @ Urbana-Champaign (`fredric3@illinois.edu`)

The wave of democratic transitions at the end of the last century held out the promise that regular elections would enable citizens to hold their leaders accountable and lead to improved governance. But many of these new democracies in the developing world are ethnically heterogeneous, returning scholars to basic questions about the compatibility of democracy with ethnic diversity (Rabushka and Shepsle, 1972; Horowitz, 1985). Competition between political parties with ethnic bases could lead to a dynamic of ethnic outbidding and a focus on turnout of core supporters, and thereby undermine and destabilize the new regimes. Such concerns were serious enough that constitutions and electoral laws frequently contain bans on ethnic parties (Basedau et al., 2007).

Many countries have sustained regular electoral competition over the last two decades without ethnic outbidding (Chandra, 2005). But elections in these new democracies have been frequently marred by electoral irregularities such as vote-buying, ballot-stuffing, and violence (Schedler, 2002) with consequences for political participation (McCann and Dominguez, 1998; Simpser, 2012) and confidence in democracy (Birch, 2008; Elklit and Reynolds, 2002; Rose and Mishler, 2009). And partly because election outcomes in diverse societies depend heavily on turnout by different ethnic groups, they have been troubled by controversies over the voters register and disputes over census results that publicly set the relative sizes of ethnic groups. Such disputes have occasionally delayed or even led to the cancellation of elections and diminished the public's acceptance of the results. Elections in diverse societies need not be "ethnic censuses" (Ferree, 2010). But this possibility incentivizes parties to engage in not only voter mobilization for turnout, but also the registration and possibly extra fraudulent registrations of core supporters to facilitate ballot box stuffing or multiple voting on election day. The updated question for new democracies is then: how does ethnic diversity affect political parties' strategies for voter registration fraud?

Although ethnicity has been central to the scholarship on voter behavior and clientelism in new democracies and particularly in sub-Saharan Africa (Posner, 2005; Wantchekon, 2003), it has been far less prominent in studies of electoral fraud and other illegal strategies of political parties. For example, Collier and Vicente (2012) analyzes a formal model of electoral competition in which an incumbent and challenger may engage in vote-buying, intimidation, and ballot fraud (vote miscounting). The case studies note that the choice of strategies may be influenced by the strong sub-national (ethnic) identities in sub-Saharan Africa, but ethnicity is not part of the model. A separate literature comprises field or "natural" experiments that study interventions such as election observers and anti-violence campaigns to improve election quality and political participation (Asunka et al., 2013; Brancati, 2012; Collier and Vicente, 2011; Driscoll and Hidalgo, 2009; Hyde, 2007; Sjoberg, 2012). In their focus on estimating the average effect of an intervention, they may overlook the underlying model of politics on which there has been an intervention. This is both a missed opportunity to learn about the politics, as well as a potential problem where there may be spillover effects or interference between units.

This paper investigates how ethnicity affected political party strategies for fraud in the specific context of voter registration ahead of the 2008 Ghanaian elections, which were anticipated to be, and transpired to be, extremely close. Indeed, the final vote margin in the presidential election was less than 50,000 out of over 9 million votes cast. We adopt the framework of Bowers, Fredrickson and Panagopoulos (2013) to re-examine Ichino and Schündeln (2012), a field experiment with the Coalition of Domestic Election Observers (CODEO) which sought to prevent the inflation of the voters register. The concern was that because Ghanaian elections are closely observed and appear to be fairly clean on election day, much of the fraud had moved to the less scrutinized voter registration stage. The political parties could add pre-marked ballots to ballot boxes or have voters vote several times to skew the result in its favor without creating a suspiciously high turnout rate if the number of registered voters were high.

Ichino and Schündeln (2012) randomized the allocation of pre-election domestic observers during the 13-day voter registration exercise leading up to the 2008 general elections in Ghana. It finds that registration centers assigned observers had, on average, a smaller increase in voter registration from 2004 than those registration centers not assigned observers. Furthermore, it finds that registration centers without observers located near a center that was assigned an observer experienced a larger increase in registrations than would have been expected without the domestic observer intervention. It proposes that this may be explained by observers *displacing*, not just deterring, political party agents who organize false registrations. Ichino and Schündeln (2012) anticipated the possibility of these spillovers in the design of the field experiment. Lacking direct observations of these illicit activities and data on ethnicity, questions remain about the parties' strategies, the extent of false registrations, how ethnicity affects these strategies, and how observers interfered with their operations. But the experimental design and additional data on ethnicity and road networks can help us address these additional questions.

Building on theories of instrumental ethnic voting (Bates, 1983; Chandra, 2004; Ichino and Nathan, 2013; Posner, 2005), we propose a model of political party strategy for inflating the voters register that anticipates citizen reactions to this form of fraud. The basic idea is that citizens who strongly prefer that a particular party win the election may be tacitly complicit when his favored party attempts to add false registrants to the voters register. Anticipating this, and using ethnicity as an indicator of who is likely to be supportive of the party, a political party's agents target registration centers in areas dominated by its associated ethnic groups. Moreover, registration centers are embedded in road networks, allowing agents to travel from one accommodating registration center to another. When party agents encounter an observer at a registration center, they will divert their illicit activities to nearby registration centers located in areas with similarly favorable ethnic demographics.

This theory is formalized in a model, along with a simpler model in which party agents move to the nearest registration center along a road network without consideration of ethnic demographics.

We then characterize support for the model in the framework for statistical inferences for causal models proposed by Bowers, Fredrickson and Panagopoulos (2013). As we explain below, our theoretical models and statistical assessments here are useful but unorthodox. We provide neither a likelihood function nor comparative statics, but still use data to learn about a model.

Our preliminary results support the idea that both road distance from a treated unit (and distance from other registration centers) and ethnic homogeneity inflate false registrations — although when the total number of false registrations implied by the models and rendered not-implausible by the data is divided equally by the number of registration stations, the effect appears small.

The paper proceeds as follows. We first verbally describe our model of how ethnicity might affect the behavior of political party agents. Section 2 provides background information on the 2008 Ghanaian elections and voter registration exercise, details of the field experiment, and the network data. We then formalize the verbal model and elaborate on the framework for assessing models with network treatment effects using the design and data from the field experiment. Analysis and discussion of the models follow.

1 Ethnicity and Registration Rigging

How do registration observers affect inflation of the voters' register? Ichino and Schündeln (2012) specify simple incentives for the political parties. Because inflating the voters register with false registrations is illegal, and if discovered, could trigger a revision of the register that undoes their work, the party's agents prefer to avoid civil society observers who will call attention to their activities. Party agents may move their activities to another nearby registration center if they encounter an observer at a registration center. Observers might therefore deter some false registrations at the registration centers at which they are stationed, but also divert the illicit activities to nearby registration centers.

This simple model did not feature differences among ordinary voters and their interests, and how these might affect the strategies of the political parties. But in high stakes elections in new democracies in ethnically divided societies, citizens often believe that politicians reward their co-ethnics or core supporters, and hence prefer to have their own co-ethnic in office. Candidates also encourage this view, along with the implication that other groups are supporting their own candidates (Bates, 1983; Posner, 2005). Politicians may also stoke fears that having a non-coethnic in office can diminish voters' livelihoods and security, resulting in a "winner-take-all" type politics in which supporting an (incumbent) politician from one's own group is the best option for a voter (Padro i Miquel, 2007). Indeed, many political parties have strong ethnic profiles, and voters use ethnicity as an informational shortcut for what constituency will be served and whose clientelistic promises are more credible (Conroy-Krutz, 2013; Wantchekon, 2003). Voters who strongly prefer that one party win over another may be willing to help it win the election by ignoring and therefore tacitly supporting false registrations by their own party.

Moreover, citizens in ethnically homogeneous areas or "homeland" areas of ethnic groups associated with a particular political party may be more willing than citizens in mixed areas to be complicit with false registrations. Because many of the benefits provided by government in rural areas are club or local public goods that can be targeted to particular communities (Ichino and Nathan, 2013), citizens in more homogeneous areas have more at stake in the election than their counterparts in areas without many people from the ethnic group(s) affiliated with a major party.

Anticipating that citizens who want the political party associated with their ethnic group to win will be complicitous with its illicit activities, political parties target registration centers in more ethnically homogeneous areas for false registrations. And when their agents encounter observers at those registration centers, they redirect their activities to nearby, similarly ethnically homogeneous and accommodating registration centers.

However, political parties may not favor registration centers in more ethnically homogeneous areas, if ordinary citizens would generally do nothing upon seeing a small crowd of unfamiliar people registering to vote. Ordinary citizens may not expect to recognize other people at rural registration centers that serve multiple villages in ethnically diverse areas. Even if they thought that the small crowd of registrants was suspicious, they may not want to challenge this crowd and possibly invite later trouble with the political parties. In this case, the party agents would start out at registration centers that are located near one another. If they encounter an observer, they would move to other nearby registration centers, without regard for the ethnic composition of the areas around those centers. Then ethnicity would play little role in the parties' strategies for registration fraud.

After describing the experiment and the 2008 Ghanaian elections in the next section, we formalize this verbal model. From these micro-foundations, we construct the registration figures we would see at each registration center under different causal models, which we then assess against the observed data.

2 A Field Experiment on Voter Registration in Ghana 2008

2.1 The 2008 General Elections in Ghana

Ghana, an ethnically diverse country of about 23 million in West Africa, has held regular, competitive elections every four years since its transition to democracy in 1992. The independent Electoral Commission of Ghana organizes concurrent, direct elections for president and a unicameral national parliament. The parliament is composed of 230 members elected by plurality from single-member districts. A president may serve only two terms and is elected by majority in a single national district; should a candidate fail to win a majority in the first round, the top two finishers compete in a run-off election.

Election day activities in Ghana are carefully monitored by well-organized domestic groups and international observer missions. But voter registration, which takes place several months before the election, is not routinely or broadly monitored, so both civil society groups and ordinary citizens widely suspect that the voters register is full of false registrations that can be used to manipulate the election outcome. The National Democratic Convention (NDC) and the then incumbent New Patriotic Party (NPP), the two main political parties, regularly accuse each other of inflating the voters register to their own advantage, making voter registration a major point of contention and a threat to the integrity and legitimacy of the election outcome (Ichino and Schündeln, 2012).

Citizens of Ghana may only register to vote during designated registration periods. They must register in person and only at the registration centers associated with their polling station which is tied to their residence. Each electoral area is composed of several polling stations, and in practice, the voter registration center is located at one of the larger polling stations in the electoral area. Because Ghana does not have a national ID system, it is fairly easy to declare false information. The penalty for giving false information or registering multiple times is up to a year in prison, but very few people are ever prosecuted for false registration. The voter ID card may have a photograph if a camera was available at the registration site. On election day, a voter must go to the particular polling station associated with his residence and present his voter ID card in order to vote. Ichino and Schündeln (2012) describes the voter registration process for 2008 in greater detail. Partly because of the problems associated with voter registration exercise in 2008, Ghana adopted a new system of biometric voter registration for the 2012 elections.

For a variety of reasons, including shortage of equipment and controversy around suspicious high registration figures in the 2006 voters register, the 2008 voter registration period was delayed several times. It finally began on 31 July 2008, with only a day's notice. Although each electoral area was supposed to have its own registration workstation, only about 2500 workstations were available for approximately 4800 electoral areas. Workstation distribution appeared to be disorganized, and no advance information on which electoral areas would have workstations on what days was centrally available. On the last day of the registration period, the Electoral Commission extended this period from 11 days to 13 days, due to shortages of materials and equipment.

Although the Electoral Commission had estimated that 800,000 people had become newly eligible to vote since the last voter registration period, there were approximately 2 million new registrations on the provisional new register as of the end of the 2008 registration period. As in previous voter registration exercises, registration observers and others saw political parties transport people to registration centers on buses and trucks (Ichino and Schündeln, 2012). After vetting for deceased voters and other processing by the Electoral Commission, the final voters register listed approximately 12.5 million voters.

The general elections were held without delay on 7 December 2008, and as expected, the 2008 elections were extremely competitive. Both major parties performed very well in the areas historically identified as their respective regional and ethnic bases. The opposition NDC won the presidency with a final official vote margin of less than 50,000 votes in the runoff election.

2.2 Experimental Design

The field experiment was conducted in cooperation with the Coalition of Domestic Election Observers (CODEO), an umbrella group of Ghanaian civil society organizations that monitors election campaigns and organizes observers. Before randomizing, CODEO and the researchers decided to focus the experiment on four of the ten regions of Ghana: Ashanti, Brong Ahafo, Greater Accra, and Northern Regions. These four regions contain approximately 54% of the Ghanaian population as of the 2000 Census, containing 116 of its 230 parliamentary constituencies and 2,204 of its approximately 4,800 electoral areas (ELAs, which are subunits of constituencies). Selection of regions was not random and reflected both practical concerns about the availability of observers, budget, and a desire to include as much of the population of Ghana as possible given those constraints.

The constituencies were organized into between two to four blocks in each region by the difference in vote share won by the NPP and NDC candidates in the 2004 parliamentary elections. This vote margin could be as great as 69 percentage points. In the first stage of the randomization, we randomly selected from each block two constituencies to have no observers (control constituencies) and one to have observers (treatment constituency). In the second stage, in each of the treatment constituencies, we randomly selected approximately 25% of the electoral areas to be visited by registration observers. This resulted in 77 ELAs assigned to the treatment condition and 791 ELAs to the control condition. The outcome is the number of registered voters, which was obtained from the Electoral Commission of Ghana. To control for variations in the size of the ELA, we also obtained registration figures for the 2004 election and subtracted these from the 2008 registration counts, making the final outcome the difference in registration across these two elections.

The observers were recruited from CODEO member organizations through their usual procedures, and were trained by CODEO and the research team in Accra at the offices of the Ghana Center for Democratic Development (CDD-Ghana), which serves as CODEO's secretariat. They deployed as observers with official accreditation from the Electoral Commission. They were instructed to visit unannounced only the registration centers for the ELAs on their list, to make an initial visit of 1–2 hours and then return for up to a full day in a later visit on an unannounced date. Observers in Ghana may not assist or interfere in voter registration, but they may interact with the officials and party agents present. The observers faxed or e-mailed in their checklist for each registration center every couple days, and these checklists were collated by CODEO.

2.3 Data

We added road network and ethnicity data to the data on treatment assignments and official voter registration figures used in Ichino and Schündeln (2012). We produced the road network data by digitizing regional maps of roads and villages from the Department of Feeder Roads of the Ghana Ministry of Roads and Highways in Accra. Because most urban roads are not under the purview of this department, we drop Greater Accra Region which contains the capital city Accra from the analysis. This leaves 797 ELAs in 33 constituencies organized into 11 blocks.

These maps and the updated GNS gazetteer of place names were also used to make corrections to some of the geocoding of the electoral areas in the original study.[1] The electoral areas are each composed of several polling stations, and because no maps of the electoral area boundaries are available, the polling station in the electoral area with the greatest number of registered voters in 2004 was used to geocode the electoral areas as points. These electoral areas could then be associated with census enumeration areas, which allowed them to be assigned the ethnicity data for those enumeration areas from the 2000 Ghana Population and Housing Census from the Ghana Statistical Service. Unfortunately, ethnicity data is only available at the level of the larger ethnic groups and not their subgroups recognized by the Ghanaian government, and this may obscure some important differences across registration centers. In some cases, multiple electoral areas are coded to the same enumeration area, but in many cases the catchment area for the electoral area is larger than just the enumeration area to which the electoral area is coded.

The distance along the road network from one electoral area to another was calculated as the sum of the direct distance from the origin electoral area to its closest point on the network, the direct distance from the destination electoral area to its closest point on the network, and the distance along the network between those points. The remaining data on observer assignments and the number of registered voters at the ELA-level come from the original analysis of the field experiment.

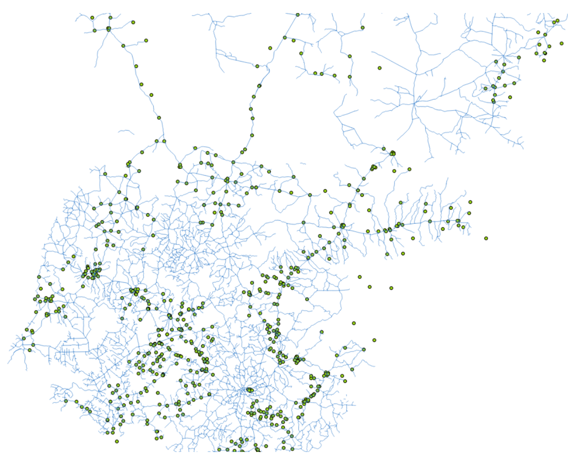

Figure 1: A portion of the road network connecting electoral areas in Ghana.

[1]The GNS data are available from the United States National Geospatial-Intelligence Agency (NGA) at `earth-info.nga.mil/gns/html/`.

2.4 Registered voters in treatment and control ELA

The left panel of Figure 2 shows the outcomes from the study centered by constituency and block. On this outcome, ELAs without an observer had a median change in the number of registered voters of 40, while ELAs that were assigned an observer had a median change in the number of registered voters of 103. Among the control ELAs, 40% had more registered voters in 2008 than they did in 2004, as compared with 26% of the treated ELAs. The treated and control ELAs differ in their distribution of the outcomes, as can be seen in the left panel of Figure 2.

This figure presents our questions: how much of this difference in distributions can be attributed to the observers? Moreover, if observers were able to prevent party agents from inflating voter roles, how much of that activity was redirected to nearby registration centers? In other words, how much of the increase in the number of registered voters in the control units would we have even in the absence of the field experiment? More importantly, did the presence of a party's ethnic base enhance the effect of observers in certain treated ELAs? If party agents indeed diverted their fraudulent activities elsewhere, did they target areas with concentrations of co-ethnics who are likely to be partisan supporters?

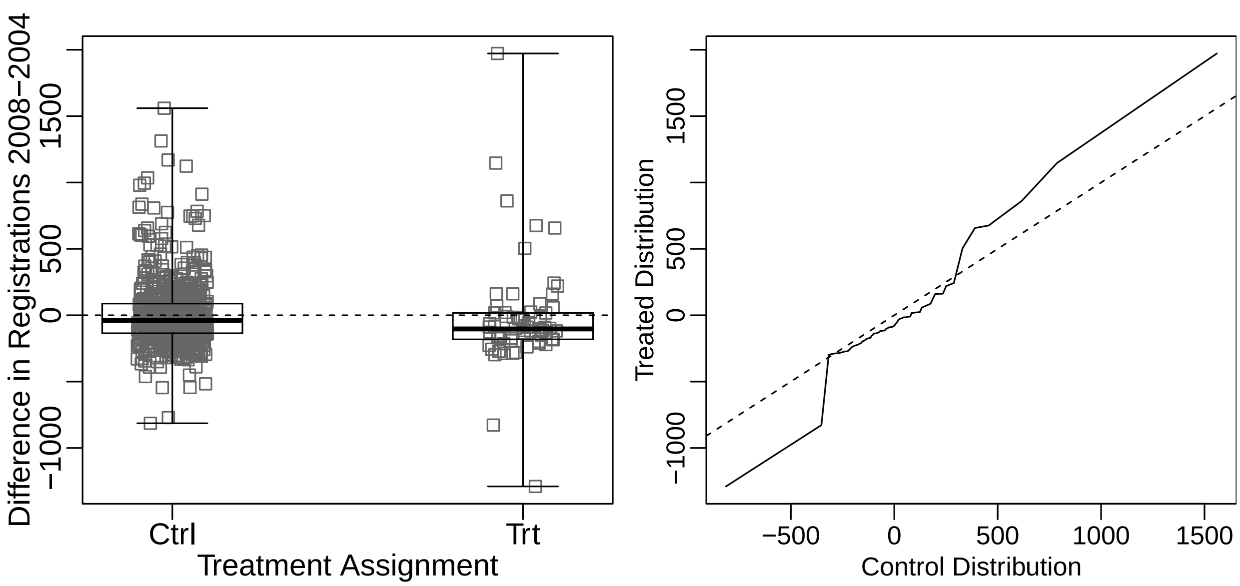

Figure 2: Left: The number of registered voters was lower in 2008 than in 2004 in treated ELAs (centered within experimental block). Right: QQ-plot of change in the number of registered voters in control and treated ELAs. The distributions differ largely at the tails.

In the next section we present models that assess the plausibility of explanations accounting for the patterns observed in Figure 2. We premise our analysis on the maxim, "set a thief to catch a thief." If party agents are subverting the voter registration process, we can only understand these outcomes by understanding the motivations and constraints faced by party agents. Our models start by considering the goals of party agents and how they would achieve those objectives. From these micro-foundations, we derive the overall impact these agents would have on voter registration counts in each ELA. Using these implications, we test hypotheses that capture the size and scope of corrupt activity. In the end, the goal of these models is to remove the patterns we see in the boxplot and QQ-plot of Figure 2. If our models are good approximations to the actual activities of party agents, we can make the two boxplots in Figure 2 (left) would look like two random samples from the same population, rather than looking as different as they do now. In the next section we provide the details of these models and the statistical methods we use to evaluate them.

3 Formalizing and Assessing Models of Network Treatment Effects

A scientific[2] model makes a statement about each observable unit in a study.[3] A causal model makes this statement in counterfactual form. For example, a very simple model might say that, had ELA i been assigned to treatment it would have registered $\tau = 20$ more people if it had been in the control condition because the buses used to transport party activists between polling stations held approximately 20 people and only one bus was deterred from unloading at this ELA by observers. We can formalize this counterfactual model as $y_{i,Z_i=0,\mathbf{z}_{-i}} = y_{i,Z_i=1,\mathbf{z}_{-i}} + \tau$, where $y_{i,Z_i=0,\mathbf{z}_{-i}}$ is the outcome we would have seen for unit i if it had been assigned to the control group ($Z_i = 0$) and the other units in the study ($-i$) were assigned treatment according to some vector of treatments $\mathbf{z}_{-i}$. We often write the potential outcome for unit i in the control condition as $y_{i,Z_i=0,\mathbf{z}_{-i}} = y_{i,Z_i=0,\mathbf{z}'_{-i}} = y_{i,Z_i=0} \equiv y_{i,0}$ where $\mathbf{z}_{-i} \neq \mathbf{z}'_{-i}$.[4] This means that the outcome that we would have observed for unit i in the control condition would be the same regardless of the treatments assigned to the other units in the study. Of course, we can, and will, relax this equivalence using a model.

This first model is a simple model of constant additive effects based on an observation from the field. Later, we will show how to assess whether there are any values of model parameters, like τ above, that are implausible given our data. While we cannot say for certain that social actors

[2]We say "scientific" here to distinguish our model from a model of a data generating process — like a model of means or a model of a stochastic process. Here, the idea is that a scientific theory would have something specific to say about all the units in its domain. That is, a physical theory would have something to say about any given rock falling under certain conditions. Perhaps the readers can help us with a better word? We do not mean to imply that other models are not serious, transparent, replicable, and capable of leading to cumulation of knowledge.

[3]It also may make a statement about units not in a study, but the relationship between a model and units not observed (in times and places unknown) is beyond the scope of this paper.

[4]The idea that we might conceptualize and formalize causal relations for experiments using counterfactuals stems from Neyman (1923 [1990]) although the development and formalization of the idea of potential outcomes is mostly traced to Rubin (1974). Excellent expositions of these and related ideas for social scientists may be found in Brady (2008); Sekhon (2008).

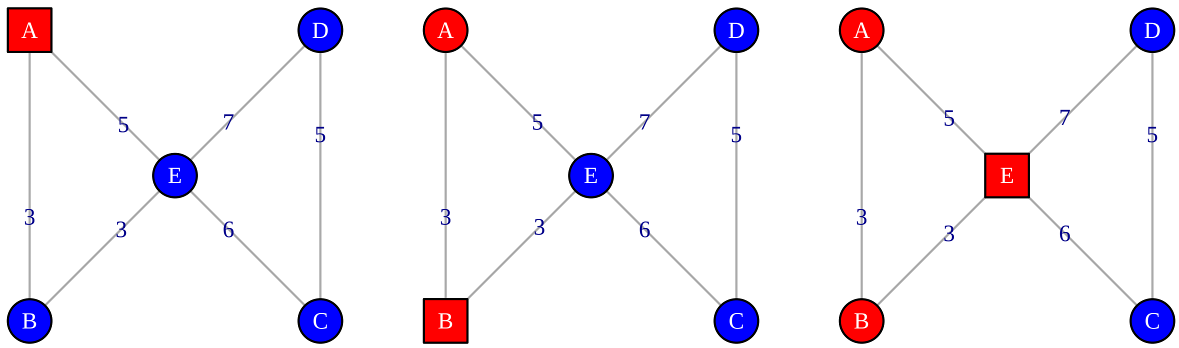

Figure 3: Agent movement rules when no observers are encountered. Squares indicate the agent's current location. Red ELAs are visited. Blue ELAs are not yet visited. From left to right: 1) $t = 0$, the agent starts at A, 2) agent selects B as closest ELA, 3) Agent moves to E in final period.

behave according to the agents in the model, we will show that sets of rules and environments that *do not* lead to observed social patterns can be discarded. A model that is not substantively incommensurate with a set of data and hence not discarded is a *candidate explanation* for a social phenomenon (Epstein and Axtell, 1996; Epstein, 2006).

In our causal scientific model, political party agents want to add false registrations to the voters register during the pre-election voter registration period to facilitate and cover up ballot stuffing or multiple voting at the general election. For now, we do not differentiate agents of different parties, and all agents share the same behavioral rules and abilities. At the beginning of the registration period we imagine that a total of k agents are placed at voter registration centers (ELAs). The party agents know their own location and the distances to all other ELAs. Agents may then to move from ELA to ELA within a constituency for a set number of iterations t, which we call "ticks." At each ELA an agent visits, it register by a given amount, τ, if an observer is not present. If an observer is present, the agent may immediately move to a new ELA or stop traveling if no other ELAs are more attractive (we explain how different models of behavior make different ELAs differentially attractive in a couple of paragraphs). An agent may not visit an ELA that he has already visited, but multiple agents may visit the same ELA. The observers randomized to ELAs to prevent registration fraud do not move, and their locations correspond to the location of observers in the fielded study. The observers are fixed features of the environment, and the only actors in the model are the agents.

Figure 3 illustrates these rules governing agent movement for five ELAs (A, B, C, D, and E) connected by roads. In this example, there are no observers present and we presume a single constituency. The figures show a party agent starting at ELA A and moving through the network for two ticks. The agent's position is marked with a square. Visited ELAs are red, while unvisited ELAs are marked in blue. The distances between ELAs are marked on the roads between them. Although not marked, the agent could travel from A or B to C or D in a single tick by passing through E without stopping. However, for this simple example, the agent will always prefer the shorter, single edge paths. At time zero, the agent starts at A, and considers its options of distances: $(B = 3, C = 11, D = 12, E = 5)$. Since B has the smallest value, it moves to B in tick 1. The agent then reviews the options: $(C = 9, D = 10, E = 3)$. As E is the minimum distance, the agent moves to E, and the simulation ends.

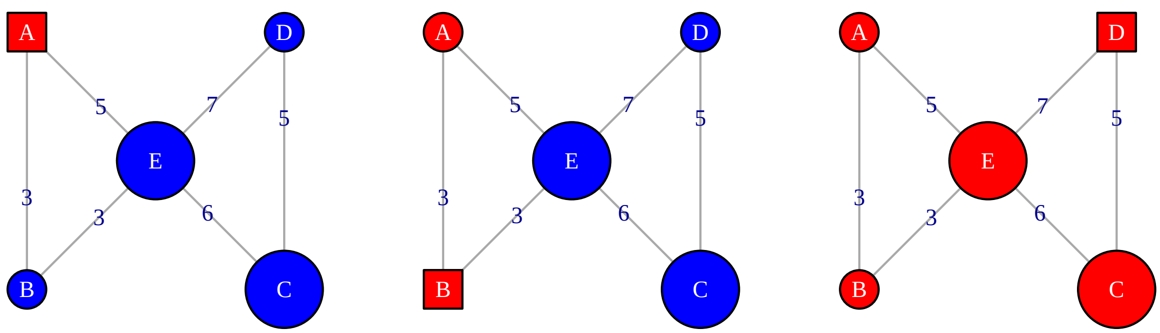

Figure 4: Agent movement rules when observers are present. Squares indicate the agent's current location. Red ELAs are visited. Blue ELAs are not yet visited. The large circles indicate observer ELAs. From left to right: 1) $t = 0$, the agent starts at A, 2) agent selects B as closest ELA, 3) Agent moves to E, but as an observer is present, immediately moves to C, again encounters an observer, and finally stops at D.

Figure 4 shows the party agent's movements when observers are placed at ELAs C and E, indicated by large circles. The agent does not know that there are observers at these locations in advance. In the first time period, the agent makes the same choice as before and moves to B. In the second period, the agent again chooses E, but because an observer is located at E, the agent immediately moves to a new location. As A and B have already been visited the agent selects between $(C = 6, D = 7)$. It chooses C, but again finds an observer and must go to D. As D does not have an observer, the agent stops there, ending the simulation.

To summarize, this framework has five parameters:

k The total number of agents.

t The total number of "ticks," or time periods, in which agents can visit ELAs.

τ The number of false registrants an agent can add to an unobserved ELA.

Movement rule The rule by which agents move to their next target. If an agent cannot move to another ELA, he becomes inactive for the remainder of the registration period.

Starting location rule For now, we imagine that the agents can rank all of the ELAs according to how easy it would be to inflate the register in those places. If they can do this, then we posit that they will choose the most desirable ELA as their starting positions.

This raises the following question: what makes a registration center (ELA) attractive to party agents? The theoretical discussion above suggests and distinguishes between two different models of party agent behavior – one that explicitly considers the ethnic composition of the areas around a registration center and another that only considers travel distance. We will show that these models imply different patterns of outcomes for the ELAs. Therefore, our observed data will allow us to assess whether and how each of the models may be a candidate explanation for registration rigging, and extend our understanding of how ethnicity affects political party behavior in new democracies.

3.1 Geographical Distance Minimizing Model

In this model, agents only wish to minimize the distance travelled from one ELA to another. Agents must visit t locations, but they try to do so by taking short steps. In this model, agents are not optimizing their entire route. At each tick, the agent considers all neighboring ELAs given their current location in the network and selects the closest. If the ELA contains an observer, the agent chooses a new location by the same decision rule. This extra move counts as part of the same round or "tick." This greedy approach may result in longer routes than would be available if the agent considered all possible routes of size t from a given starting location.

Since the party agents wish to minimize distance in one step increments, we select starting locations based on the closeness to other ELAs. That is, the most desirable ELAs are those from which it would easiest to travel to other ELAs. ELAs are ranked by their minimum distance neighbor, and the first k are selected as starting locations for the agents.[5]

[5] Right now this means that some constituencies may be allocated no agents if the number of agents is small. We may change this and also add information about the competitiveness of the elections in a given constituency in the next round.

3.2 Ethnic Homogeneity Preference Model

Like the distance minimizing model, party agents in this model have an eye towards to keeping their total travel distance small but have the additional desire to visit ethnically homogeneous areas. For each ELA, we compute the ethnic compositions over 8 groups and a residual category. The ethnic group categories are Akan, Ga-Dangbe, Ewe, Guan, Gurma, Mole-Dagbon, Grusi, and Mande. Each ELA is then given an ethnic fractionalization score based on the percentages (p) of each ethnic group:

$$F = 1 - \sum_{i=1}^{9} p_i^2 \tag{1}$$

The value of F is 0 when there is only one ethnic group in an area and 0.89 when all 9 ethnic groups (including the residual category) are equally balanced. In this simulation, agents prefer to visit ELAs that have values of F close to zero. At the beginning of the registration period, the agents rank the ELAs by F rather than by centrality in a dense road network, and k agents go to the k ELAs with the smallest values of F.

At each step of the process, agents only consider moving to an ELA with F at or below the α percentile of F in the constituency, the threshold above which an ELA becomes too ethnically heterogeneous to be worth visiting. Among these, an agent picks the closest ELA in terms of road distance. If no ELAs have sufficiently low values of F, the agent is stuck and stops his efforts to add false registrants. As before, if an agent moves to an ELA with an observer, he immediately moves, choosing the closest location with F at the α percentile or lower for the constituency. The same α percentile applies to all constituencies, but because constituencies differ in their distribution of F, the threshold in terms of ethnic homogeneity differs by constituency. We presume that agents focus their activities within particular constituencies and that therefore they are assessing the attractiveness of ELAs relative to other ELAs in their constituency.

3.3 Summary

We have two theoretical models of party agent behavior. Each model would imply different numbers of visits by different numbers of agents directing different sized vehicles to each ELA in the experimental pool. Neither model will completely describe the behavior of the system. However, evidence against either model can teach us about the larger questions on how ethnicity affects political party strategies for electoral fraud.

4 Assessing the evidence against a model

Following Bowers, Fredrickson and Panagopoulos (2013), we assess the empirical evidence against our model within a framework of Fisherian hypothesis testing. We first describe the general approach with a simple example, before applying the method to the more complicated model of political party agents.

4.1 An overview of Fisher's sharp hypothesis testing

In Fisher's original framework for hypothesis testing, a hypothesis is a statement about a counterfactual for each unit in the study. The most famous such hypothesis is the sharp null hypothesis of no effects, which we might write in terms of potential outcomes following Rosenbaum (2010, Chapter 2) as $H_0 : y_{i,1} = y_{i,0}$. Notice that this hypothesis is very similar to our model of constant additive effects ($y_{i,1} = y_{i,0} + \tau$), which only differs because we did not presume that $\tau = 0$. Bowers, Fredrickson and Panagopoulos (2013) point out that, in general, any model which generates a counter-factual statement for all units in a study can be understood to be, in essence, a sharp hypothesis generator. We can therefore engage such models using very well established procedures for hypothesis testing in a randomized study.[6]

Recall our simple model of constant, additive effects for a unit i: $y_{i,1} = y_{i,0} + \tau$. This model encodes not only a hunch about how the treatment operated (here, via simply adding τ to all the potential outcomes under control), but also a statement that the potential outcomes of unit i do not depend on the potential outcomes of other units. What does this model imply if τ were 100? Figure 5 shows one vision of what this model implies. On the left hand side we see the outcomes in the control and treated groups compared just as they were in Figure 2. The right hand side shows the two outcome distributions after removing $\tau = 100$ from the treated group. If the simple model described the data well, then removing 100 from each and every unit in the treated group would make the treated and control groups look like they were random samples drawn from the same population, the set of ELAs in the study. Note that this model shifted the treated group distribution down, but it did not appreciably change the distribution of the treated group to make it look especially similar to the control group. The graphical evidence here suggests that our observed data (shown at left) would be more surprising from the perspective of hypothesis of $\tau = 100$ given the model $y_{i,1} = y_{i,0} + \tau$ than would be the hypothesis that $\tau = 0$ (i.e., the sharp null hypothesis of no effects). In fact, when we compare our observed data to what would be implied by each hypothesis, $\tau = 0$ and $\tau = 100$, we learn that it would be quite surprising to see our observed data from the perspective of either hypothesis, with $p = 0.027$ for $H_0 : \tau = 0$ and $p = 0$ for $H_0 : \tau = 100$.[7] Notice that the low p-values refer to both the model and the particular parameters of the model and the test statistic: the test statistic provides enough power against relevant alternatives, and our

[6] See Rosenbaum (2010, Chapter 2) for a clear and modern description of Fisher's approach to statistical inference. Of course, Fisher's own "Lady Tasting Tea" example is also highly recommended (Fisher, 1935, Chapter 2).

[7] We are glossing over the choice of test statistic for this example. Here we use the KS test statistic which gauges the difference between two distributions using the maximum distance between the ECDFs of two distributions. We repeated the two stage sampling process (in blocks, etc..) 500 times and recorded the proportion of the distribution of the distribution representing the hypothesis greater than or equal than our observed value as the p-value. Thus, we make no large sample assumptions here.

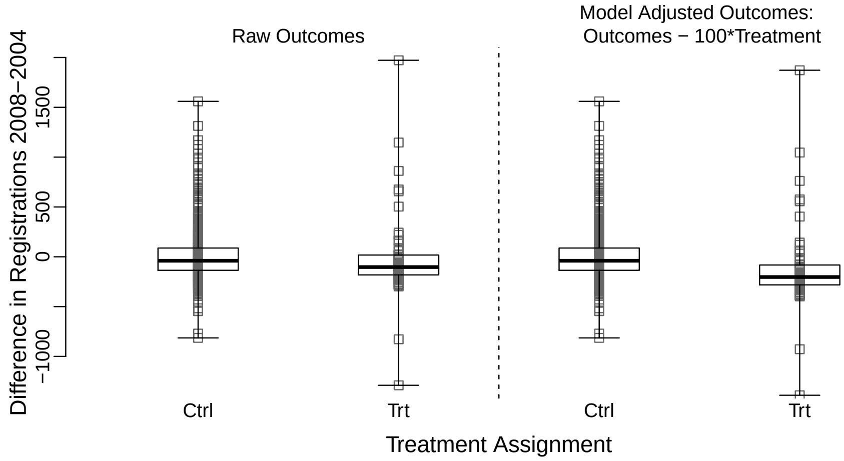

Figure 5: An example of outcomes implied by a model. Left boxplots show original outcomes. Right boxplots show the outcomes implied by $y_{i,1} = y_{i,0} + \tau$, $\tau = 100$.

model is sensitive enough to changes in τ, to distinguish the two groups under our two hypotheses.

4.2 Assessing sharp models of party agents

Now let us return to our models of party agents engaged in inflating the voters register. Recall that our models are broadly distinguished from each other by movement and starting rules that consider distance only or also consider the ethnic homogeneity of the area, and that for a given selection of τ, k, t, our models generate a specific number of false registrations for each ELA, V_i, which we collect into a vector $\mathbf{V}$. To assess the evidence against our models, we select a range of values for k, t, and τ, as shown in Table 1. We presume that k, the number of party agents, is limited by the number of total ELAs: if one party agent began with a bus load of party activists at every single ELA we would have expected more widespread reports of attempts at registration rigging. Then for each model and set of parameters, we run two simulations. In the first simulation, we model the agents' behavior in the observed experiment in which observers are placed at the ELAs assigned to treatment. The agents are placed in their starting locations and proceed to visit ELAs according to the decision rules in the model. If an agent encounters an observer, he immediately moves to another location within the same "tick" period. Once the simulation has worked through the iterations, we have a count of the number of times each ELA was visited by agents for that model and parameter set.

Then we remove all the observers and repeat the simulation. This second simulation corresponds to the world in which the experiment was never run (or the "uniformity trial" (Rosenbaum, 2007)). Agents start at the same location, but because there are no observers, they will stop at one ELA per tick and register false names. At the end of this simulation we again have a list of how many times each ELA was visited for this world with no observers.

	Minimum	Maximum
k	0.00	797.00
ticks	1.00	20.00
tau	-25.00	100.00
alpha	0.00	1.00

Table 1: Parameter range investigated for the two models. We presume that k, the number of party agents, is limited by the number of total ELAs. t is the number of iterations of the simulation or number of ELAs at which registration rigging would have been attempted. τ is the number of false registrations each party agent can add to an ELA he visits. The quantiles of ethnic homogeneity (α) that determine the attractiveness of a place under the ethnicity model vary by constituency.

Using these two sets of counts, which we call $\mathbf{V_z}$ and $\mathbf{V_0}$, respectively, for the experimental and "uniformity trial" with no observers, we can adjust the observed outcome to the outcome that would have been observed if no experiment had been conducted. If our models isomorphic to the true process by which treatment propagated through the network, distinguishing the distributions of the treated and control groups, then the treated and control groups would be random samples from the observer-less outcome. We generate this outcome by asking, "what is the difference between the observed (experimental) data and the world in which no experiment had occurred?" In these models, if an observer were at a site during the intervention ($Z_i = 1$), agents are unable to change the number of registrations at that ELA. However, if an agent had not been present, the registration count would have been inflated by τ for each agent, or $\tau V_{\mathbf{0},i}$. Therefore, for any treated unit, its outcome without any observers present would have been $y_{i,\mathbf{0}} = y_{i,\mathbf{z}} + \tau V_{i,\mathbf{0}} \,|\, z_i = 1$.

Conversely, if an observer was not present at ELA i in the experiment, in both worlds all agents were able to register τ false names. Thus, to adjust the observed outcome to get the observer-less outcome, we care about the difference in the number of agents that visited in the experiment as compared to the world in which no observers were present. Each additional agent that visited during the experiment adds τ additional false registrations, so we can adjust the observed outcome to the unobserved one as $y_{i,\mathbf{0}} = y_{i,\mathbf{z}} - \tau(V_{i,\mathbf{z}} - V_{i,\mathbf{0}}) \,|\, z_i = 0$. We can combine this statement with that for the treated units to achieve the overall model of effects:

$$\mathbf{y_0} = \mathbf{y_z} + \tau\left(\mathbf{z}\mathbf{V_0} - (1-\mathbf{z})(\mathbf{V_z} - \mathbf{V_0})\right) = \mathbf{y_z} + \tau\left(\mathbf{V_z}(\mathbf{z}-1) + \mathbf{V_0}\right) \tag{2}$$

This model of effects can be illustrated by again referring to Figures 3 and 4. In Figure 3, which corresponds to the uniformity trial with no observers, the party agent visited (A, B, E). In the second figure, which corresponds to the experimental trial, the agent visited (A, B, E, C, D). The effect of the observers at C and E is the difference between these two paths. The agent visited both A and B in both worlds, and neither ELA had an observer. Therefore, the outcome $y_{A,\mathbf{0}} = y_{A,\mathbf{z}}$ and $y_{B,\mathbf{0}} = y_{B,\mathbf{z}}$, and no adjustment is necessary. In the uniformity trial (the world without observers), E was visited by a party agent who was able to add false registrations. In the experiment, E was again visited by a party agent, but he encountered an observer and was unable to register voters. Therefore the outcome of this ELA is $y_{E,\mathbf{0}} = y_{E,\mathbf{z}} + \tau$. The agent did not visit C in the uniformity trial, and although C was visited by the agent in the experiment, the observer prevented the addition of false registrations. Therefore, no adjustment is necessary for C. Finally, the agent did not have time to visit D in the uniformity trial, but because observers deterred activity at E and C, an agent does visit D in the experiment. Thus, D's outcome is $y_{D,\mathbf{0}} = y_{D,\mathbf{z}} - \tau$. By repeating this logic for all the agents in the model, we can adjust our observed outcome in the experiment to that which would have been observed in the uniformity trial.

We apply the methods of Bowers, Fredrickson and Panagopoulos (2013) to test the hypotheses generated by our models of party agents. In addition to the model of effects in Equation 2, this statistical approach requires that we define how the treatment was randomly allocated and provide a test statistic to score the residual differences between the treated and control groups that remain after adjusting for the hypothesized effects of the agents. The randomization procedure was a hierarchical, blocked randomization, as discussed in Section 2.2. For a test statistic, we elect to look at the KS test-statistic as described above after adjustment.[8] Our model implies a shift in the distribution of treated units in the presence of observers, but with several ELAs with very large numbers of registrations, the mean-difference may not be sensitive to models that change the shapes of the distributions more than the location. The results of these statistical tests are discussed in the next section.

5 Analysis

5.1 How surprising is the sharp null of no effects?

We start by asking, "What is the probability of seeing the observed difference between the treated and control groups, *if the observers had no effect at all?*" This particular hypothesis is frequently referred to as "the sharp null of no effect." If this hypothesis were true (i.e., observers had no effect), any difference in the observed difference in between the two groups is purely due to the chance introduced by random assignment. In fact, if we were to reassign treatment repeatedly using the same blocked design, we could exactly compute the distribution of differences that would occur just due to chance in this study, conditional on observers having no effect. By comparing the observed difference between the treated and control groups to this distribution, we have an answer to our question: if the observers truly had no effect, the probability of seeing a test-statistic of 0.2 or greater is $p = 0.018$, an event that would happen in fewer than 1 in 55 of the possible ways that treatment could have been randomized in this study.[9]

[8] We are exploring alternative test-statistics such as Neyman's smooth test of difference of distribution (Ledwina, 1994) on the relative distribution of our two groups (Handcock and Morris, 1999). So far we have not found large differences in power.

[9] There are several parameter combinations for the our models that generate the same adjustment as the sharp null (i.e., no adjustment at all). Whenever the number of agents is zero ($k = 0$), the number of ELAs visited is zero

5.2 How surprising are our models?

We now turn our attention to the question about the implications of our models. Each combination of parameters (k, τ, t, α) implies a particular amount of false registrations for each ELA. However, each combination of parameters does not imply a unique total. Because agents stop when no ELA is more attractive under either the distance or ethnic homogeneity models, and because the model itself is additive, we can see many combinations of parameters yielding the same total number of false registrations. Although we could present evidence against each unique set of parameter values, here we focus on two quantities: the percentile of the constituency specific homogeneity distribution α and the total number of false registrations prevented by the treatment $T = \sum_{i=1}^{n}(Y_i - y_{i0})$, where y_{i0} is the number of registrations implied by the model under the uniformity trial. The total number of false registrations T can be considered a composite hypothesis, and so we summarize evidence against particular totals with the minimum p-value.[10]

Figure 6 shows the evidence against different hypotheses about total false registrations for the two models. Here we show $-T$ on the x-axis: the least implausible hypotheses about T were those where there were more registrations under the uniformity trial than actually observed, so it is easier to talk about "false registrations prevented." But the only place where the evidence was other than strongly against the hypotheses was between about 0 and 2000, as shown by the line of black dots at essentially 0 and then peaking in the area just right of 0. The ethnicity model added a parameter, α, or threshold for the minimum ethnic homogeneity required for a visit. All of the thresholds produced more or less the same pattern, shown here for three different thresholds — .1 and .25 (i.e., very few ELAs become available for a visit) and .9 (i.e., nearly all ELAs are attractive except for the ones that are extremely ethnically diverse relative to others in the constituency). On the plot we can see that the p-values for the two different thresholds are substantively equivalent. We add loess curves of those points merely to smooth what we take to be the simulation-based noise in the points.

In the most concrete terms, we see two differences between the models. First, we see that we have ample power to distinguish among similar hypotheses about totals in the distance model. Only a small range of such hypotheses are plausible at any conventional level of plausibility. For example, the totals not implausible at the .05 level range from 37 to 1125 (or 37/767 to 1125/767 if we prefer to think of a "per-ELA" amount). The ethnicity model is less sensitive to differences in the total and is not particularly sensitive to differences in thresholds of ethnic homogeneity. It, like the distance model, excludes as implausible the negative totals. However, the implausibility

$(t = 0)$ or the effect an agent is zero, $(\tau = 0)$, both models imply no adjustment to the data. The models also may imply similar adjustments when these parameters take on other, non-degenerate values, but as the parameter interact with the network structure and ethnic composition of the election registration sites, such similarities are difficult to deduce.

[10]The idea that one would reject a composite hypothesis only if *all* of the constitutive atomic hypotheses are rejected is explained in the context of a different composite causal quantity (the "attributable effect") in Rosenbaum (2001, 2002).

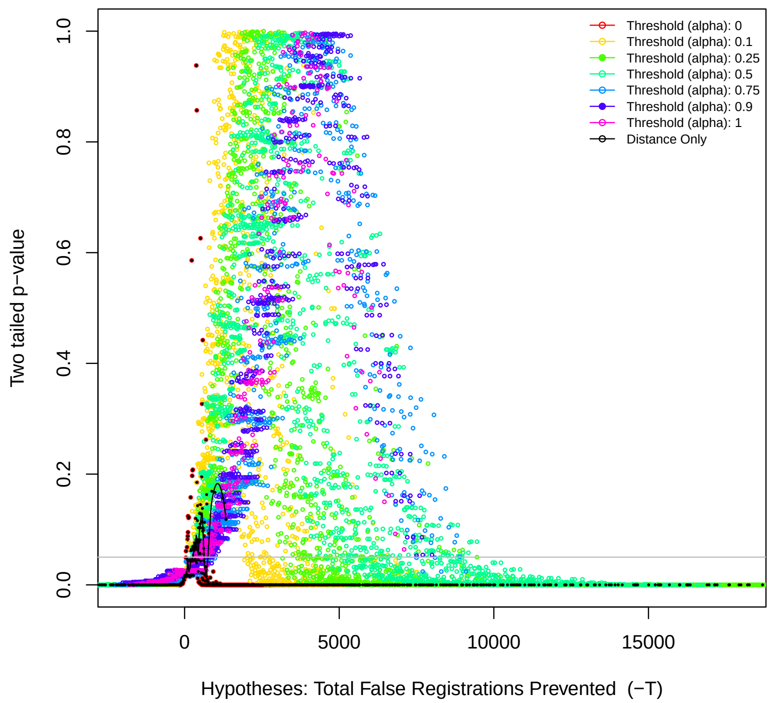

Figure 6: Evidence against hypotheses about total false registrations prevented. The open points and curves on the right reflect the model where ethnic homogeneity limits the movement of party agents for three thresholds: agents only visit ELAs more homogeneous than 90% of their constituencies, $\alpha = .1$, agents only visit ELAs in the top quartile of homogeneity in their constituencies, $\alpha = .25$ or agents only avoid the top 10% most heterogeneous places, $\alpha = .9$. The black points near 0 reflect the p-values for the distance-only model. The horizontal gray line is at $p = .05$. The curves simply smooth the points using loess.

region of the ethnicity model extends further above 0, as most of the threshold models tend to have $p = .05$ around 280. The upper bound for the implausibility of the ethnicity models is not defined here at conventional levels of surprise or plausibility for the range of parameters that we assessed for the agent-level models and the test statistic that we used to compare the resulting distributions of treated and control ELAs. However, the largest value across the thresholds tended to be around 6800 (with p-values around .1).

What do these results tell us about our models? Neither model is entirely implausible. We have seen in Bowers, Fredrickson and Panagopoulos (2013) that certain models and test statistic combinations can have no support at all from a given design and data set. We have also seen situations where the model is entirely insensitive to changes in the parameters or a test statistic such that the p-value would be near 1 for all hypotheses. In this paper, we are in a more interesting and intermediate situation: both models are implausible for certain parameter ranges, namely negative numbers of false registrations, but there are certain ranges of parameters that would not be surprising from the perspective of either model.

How would we think about two models with different support in the data? If the both models allowed for equally powerful tests, then we might consider the model with more not-implausible values as the one with most support. [We do not know right now whether these tests are equally powerful and suspect that they are not.] It is also possible for both models to be not implausible. After all, the sharp null of no effects is also implausible.

That both distance and ethnicity may be involved in the propagation of the effects of the observers across the road network is suggested in the following descriptive results where the simple differences in constituency-centered registrations 2008 versus 2004 is predicted separately for the control and treated ELAs. Figure 7 shows smoothed response surfaces for the control and treated groups. The "highest" point(s) on the surface for the control group occur either when a treated ELA is very near and the ELA itself is very homogeneous (homogeneity near 0) or when a treated ELA is distant and the ELA itself is very heterogeneous, although data are relatively sparse in that upper right corner of the plot. The treated group is much smaller (n=68 when excluding Greater Accra as we have for this paper) and the data are relatively sparse. The treated ELAs with the most positive change (i.e. most plausible registration rigging) — other than the region based only 3 observations — are in the center of the plot: places with medium ethnic homogeneity and not too distant from another treated ELA. The role that road distance plays in these displays differs from the role it plays in the agent simulations. Yet, we see that the response surface is not completely flat with respect to distance or ethnic homogeneity. Of course, these observational comparisons among control and treated units reflect many other differences between these units.

We can describe how distance and ethnicity might moderate the effect of the observers on registration inflation by comparing the two response surfaces shown above in figure 7. Figure 8 shows the result of subtracting the control surface from the treated surface. The univariate distributions of distance and ethnic homogeneity are displayed as rug plots on the side of the contour plot with the treated distribution inside the axis and the control distribution outside the axis. We see that treated ELAs tended to register more people than control ELAs when ethnic homogeneity is in the middle of the national scale and when the nearest treated ELA is fairly close (this is the yellow blob with the label "32" in the middle of the left hand of the plot).[11]

5.3 Summary

In the end, remember, that we only observe inputs (design, network location, and record of treatment administration) and output (number of registrations). Any number of complex processes may have converted the inputs into the outputs. And we do not observe those processes in any detail.

[11] We call this a "descriptive" result because we have not done the work to identify response surface as a "treatment effect" let alone discuss why this particular observed response surface might tell us something about a difference in potential response surfaces.

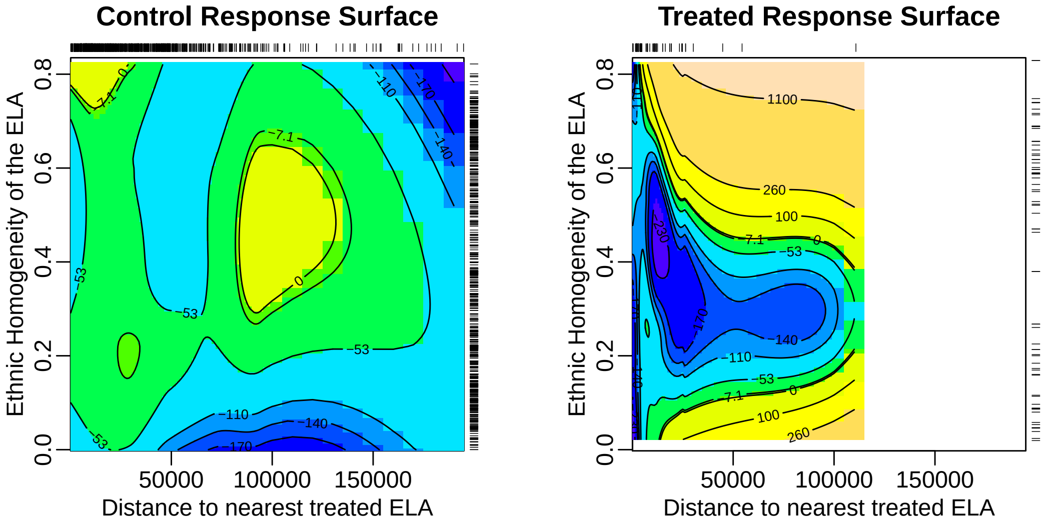

Figure 7: Response surfaces for treated and control ELAs fit with loess (span=1/2, gaussian family, degree=2) and evaluated at the same values for distance and ethnic homogeneity. Contour lines displayed at 0 and for the quantiles of the joint surfaces at c(.01,.05,.1,.2,.5,.8,.9,.95,.99).

We do have some knowledge that observers saw and it was reported to observers that buses left their registration centers heading toward other registration centers. And we also know that many more people were registered across the country, which includes more ELAs that we have here, than were eligible to do so.

We have learned about our models, however. Neither model is entirely implausible. Nor is there so little information available in our models or test statistics or data such that we cannot distinguish among parameters. Rather, if agents were to only maximize number of ELAs visited, then the total number of false registrations would be relatively small across the set of 797 ELAs. If the agents worked to maximize registration rigging by visiting possibly complicit villages rather than merely visiting as many as possible, then it is only implausible for very little registration rigging to have occurred, or if one imagines that the right hand side of the plot continued, for more than about 7000 additional registrations to have been added to the rolls (i.e., more than about 9 per ELA). The response surfaces suggest that both processes may be at work, and one might even imagine a party agent changing strategy on the fly as new information arrives by cell phone, a process that we have not modelled here.

6 Discussion and Conclusion

If party agents chose a strategy of registration rigging in Ghana in 2008 that focused on the ethnic homogeneity of geographically nearby ELAs, then it would not be surprising to see up to 9 fraudulent registrations per ELA, although it would be probable that the additional roughly 7000

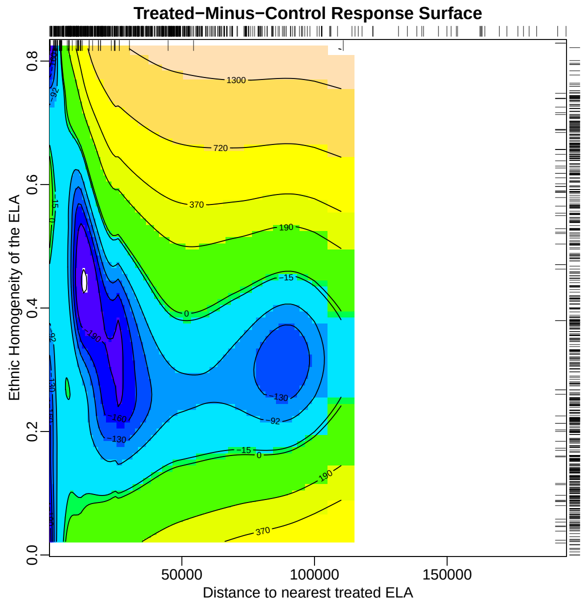

Figure 8: The difference between the treated and control response surfaces. Contours drawn at 0 and the c(.01,.05,.1,.2,.5,.8,.9,.95,.99) quantiles. Treated and control distributions of distance and ethnic homogeneity shown as rug plots — treated distribution is sparser than the control distribution and is drawn inside the axes to the plot.

registrations would not be distributed equally across all ELAs. If they focused entirely on visiting as many ELAs as possible within a constituency, then such a result would be surprising, and, in general, many fewer registrations would be counted as fraudulent (only up to about 1.5 per ELA or about 1200 in total). Some descriptive work suggests that both processes may be at work even if we have not yet specified a mixture model to assess this idea.

This work in progress combines a new way of writing models with a new way of assessing their relationship to data. The goal is to infer about a model from a randomized experiment, and our tests of the models, as shown in the previous figures, teaches us about the models, not primarily about the observations. These models involve parameters representing the unobserved movement of party agents along road networks in Ghana in response to observers at the registration stations, as well to the features of the network itself like the density of the connections and the ethnic homogeneity of the locations.

Although we need to work to build models with better parameter sensitivity, we can already exclude certain ranges of parameters and their implications from these models: that is, any model implying very large scale fraud with, say, many hundreds of agents who each manage to inflate the election rolls by many hundreds of false registrants would not be supported by the data.

Our next steps involve focusing on the model building itself, such as distinguishing between agents of the different parties – having them seek areas with the different ethnic groups that constitute their respective bases and allowing them to interact with each other – and relating the number of agents in a constituency to the competitiveness of that constituency. It also includes improving the power of our tests using better test statistics, and perhaps to delve deeper into reports of electoral irregularities surrounding this voter registration exercise in order to fine tune the details of the models.

Finally, the results that we have presented so far suggest that models where ethnicity plays a role can be useful explanations of the patterns in our data. We can never know in any detail the true mechanisms by which the treatment of observers propagates across the network. The map cannot be the terrain. However, models that we have specified so far may be useful guides to explanation for the registration outcomes in Ghana in 2008, and thus, we hope can suggest new avenues of research for how ethnicity affects the electoral process in new democracies where electoral integrity is at risk.

References

Asunka, Joseph, Sarah Brierley, Miriam Golden, Eric Kramon and George Ofosu. 2013. "Electoral Violence and Fraud in the 2012 Ghanaian Elections: Polling Station Results." Paper prepared for presentation at a Workshop on Electoral Integrity, Violence, and Vote-Buying, Center on Democracy, Development, and the Rule of Law, Program on Poverty and Governance, Stanford University, April 12-13, 2013.

Basedau, Matthias, Matthijs Bogaards, Christof Hartmann and Peter Niesen. 2007. "Ethnic Party Bans In Africa: A Research Agenda." *German Law Journal* 8(6):617–34.

Bates, Robert H. 1983. Modernization, Ethnic Competition, and the Rationality of Politics in Contemporary Africa. In *State Versus Ethnic Claims: African Policy Dilemmas*, ed. Donald Rothchild and Victor A. Olorunsola. Boulder, CO: Westview Press.

Birch, Sarah. 2008. "Electoral institutions and popular confidence in electoral processes: A cross-national analysis." *Electoral Studies* 27:305–20.

Bowers, Jake, Mark M. Fredrickson and Costas Panagopoulos. 2013. "Reasoning about Interference Between Units: A General Framework." *Political Analysis* 21(1):97–124.

Brady, Henry E. 2008. "Causation and explanation in social science." *Oxford handbook of political methodology* pp. 217–270.

Brancati, Dawn. 2012. "Building Confidence in Elections: The Case of Electoral Monitors in Kosova." Working paper, Washington University in St. Louis.

Chandra, Kanchan. 2004. *Why ethnic parties succeed: patronage and ethnic headcounts in India.* Cambridge: Cambridge University Press.

Chandra, Kanchan. 2005. "Ethnic Parties and Democratic Stability." *Perspectives on Politics* 3(2):235–252.

Collier, Paul and Pedro C. Vicente. 2011. "Votes and Violence: Experimental Evidence from a Nigerian Election." Working Paper, University of Oxford.

Collier, Paul and Pedro C. Vicente. 2012. "Violence, Bribery, and Fraud: The Political Economy of Elections in Sub-Saharan Africa." *Public Choice* 153(1-2):117–47.

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

Driscoll, Jesse and F. Daniel Hidalgo. 2009. "The Intended and Unintended Effects of Democratization Aid: Experimental Evidence from the Republic of Georgia." Presented at the 2009 Annual Meetings of the American Political Science Association.

Elklit, Jorgen and Andrew Reynolds. 2002. "The Impact of Election Adminsitration on the Legitimacy of Emerging Democracies: A New Comparative Politics Research Agenda." *Commonwealth & Comparative Politics* 40(2):86–119.

Epstein, Joshua M., ed. 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press.

Epstein, Joshua M. and Robert Axtell. 1996. *Growing artificial societies: social science from the bottom up*. Washington, DC: Brookings Institution Press.

Ferree, Karen E. 2010. *Framing the Race in South Africa: The Political Origins of Racial Census Elections*. New York: Cambridge University Press.

Fisher, R.A. 1935. *The design of experiments. 1935*. Edinburgh: Oliver and Boyd.

Handcock, Mark S. and Martina Morris. 1999. *Relative distribution methods in the Social Sciences*. New York: Springer.

Horowitz, Donald. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.

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

Ichino, Nahomi and Matthias Schündeln. 2012. "Deterring or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana." *Journal of Politics* 74(1):292–307.

Ichino, Nahomi and Noah L. Nathan. 2013. "Crossing the Line: Local Ethnic Geography and Voting in Ghana." *American Political Science Review* 107(2):344–361.

Ledwina, Teresa. 1994. "Data-driven version of Neyman's smooth test of fit." *Journal of the American Statistical Association* 89(427):1000–1005.

McCann, James A. and Jorge I. Dominguez. 1998. "Mexicans React to Electoral Fraud and Political Corruption: an Assessment of Public Opinion and Voting Behavior." *Electoral Studies* 17(4):483–503.

Neyman, J. 1923 [1990]. "On the application of probability theory to agricultural experiments. Essay on principles. Section 9 (1923)." *Statistical Science* 5:463–480. reprint. Transl. by Dabrowska and Speed.

Padro i Miquel, Gerard. 2007. "The Control of Politicians in Divided Societies: The Politics of Fear." *Review of Economic Studies* 74:1259–74.

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

Rabushka, Alvin and Kenneth Shepsle. 1972. *Politics in Plural Societies: A Theory of Democratic Instability*. Columbus: Charles Merrill.

Rose, Richard and William Mishler. 2009. "How Do Electors Respond to an "Unfair" Election? The Experience of Russians." *Post-Soviet Affairs* 25(2):118–36.

Rosenbaum, Paul R. 2001. "Effects Attributable to Treatment: Inference in Experiments and Observational Studies with a Discrete Pivot." *Biometrika* 88(1):219–231.

Rosenbaum, Paul R. 2002. "Attributing effects to treatment in matched observational studies." *Journal of the American Statistical Association* 97(457):183–192.

Rosenbaum, Paul R. 2010. *Design of Observational Studies*. Springer.
URL: *http://www.springer.com/statistics/statistical+theory+and+methods/book/978-1-4419-1212-1*

Rosenbaum, P.R. 2007. "Interference Between Units in Randomized Experiments." *Journal of the American Statistical Association* 102(477):191–200.

Rubin, D.B. 1974. "Estimating the Causal Effects of Treatments in Randomized and Nonrandomized Studies." *J. Educ. Psych.* 66:688–701.

Schedler, Andreas. 2002. "The Menu of Manipulation." *Journal of Democracy* 13(2):36–50.

Sekhon, J.S. 2008. "The Neyman-Rubin model of causal inference and estimation via matching methods." *The Oxford handbook of political methodology* pp. 271–299.

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

Sjoberg, Fredrik M. 2012. "Making Voters Count: Evidence form Field Experiments about the Efficacy of Domestic Election Observation." Harriman Institute Working Paper No 1, Columbia University.

Wantchekon, Leonard. 2003. "Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin." *World Politics* 55(3):399–422.