# Who is Favored?
# How the Outcomes One Studies Affect the Answers One Gets*

Eric Kramon
Department of Political Science
University of California, Los Angeles
kramon@ucla.edu

Daniel N. Posner
Department of Political Science
University of California, Los Angeles
dposner@polisci.ucla.edu

28 April 2010

* Research note prepared for the 30 April-1 May 2010 meeting of the Working Group in African Political Economy. We thank Brian Min for sharing his night lights data with us and for his help with various GIS aspects of our analyses.

Papers in the burgeoning empirical literature on distributive politics follow a similar pattern: the author describes a data set on the distribution of a valued patronage good; the data are explored for patterns consistent with one or another theory of distributive politics; and conclusions are drawn about the social or partisan groups that are favored in the political system and the political logic that underlies this favoritism. Thus, Diaz-Cayeros et al (2008) employ data from a major government-run anti-poverty program to look for patterns of partisan targeting in Mexico. Kasara (2007) exploits data on taxes faced by producers of different cash crops in Africa to test whether political leaders favor their ethnic kin. Golden and Picci (2008) use data on infrastructure expenditures across Italian provinces to examine legislators' and ruling parties' relative commitments to core and marginal voters. Dahlberg and Johansson (2002) employ data on government grants to municipalities in Sweden to test whether incumbents channel resources to win votes. Franck and Rainer (2009) examine data on educational attainment and infant mortality in Africa to test hypotheses about ethnic favoritism.

This literature (a sampling of recent contributions to which is provided in Table 1) is a major growth area in the field of political economy. Many of the papers within it are exceptionally creative and employ painstakingly collected data that is analyzed in sophisticated ways. However nearly all of the studies in this literature are vulnerable to a common—and potentially devastating—criticism: namely, that the pattern of favoritism identified with respect to the particular outcome(s) under study may be counterbalanced by a quite different—even opposite—pattern of favoritism with respect to other (unmeasured) outcomes. For example, districts whose roads are rehabilitated may not receive school funding; households that receive cash transfers may not get electrification; municipalities in which new clinics are constructed may not be given sanitation projects. The problem lies in the fact that governments can favor constituencies through the targeting of *multiple* public and private goods—roads, schools, clinics, electrification, cash grants and transfers, irrigation schemes, subsidies, public service jobs, and so on. Yet nearly all of the studies in the empirical distributive politics literature focus exclusively on just one (or a small subset) of these goods. So while inferences about favoritism with respect to the particular outcome being studied may be warranted, conclusions about which groups are favored *per se* rest on shaky foundations.

The purpose of this research note is to assess the severity of this potential problem. We do this by analyzing data from the standard Demographic and Health Surveys (DHS) in four African countries (Ghana, Kenya, Malawi and Zambia), taking advantage of the fact that DHS surveys collect information about multiple outcomes that are both highly valued and plausibly products of government allocation decisions.[1] We focus in particular on four outcomes: childhood vaccinations, educational attainment, access to improved water sources, and household electrification. We supplement these with a fifth measure, generated from satellite data on lights visible from space at night, on whether or not households are located in a lit cluster. This measure serves as an alternative to the data on household electrification, as gleaned from the DHS surveys. Our strategy is to compare the distribution of these five outcomes across households with differing political partisanship (defined by whether or not the household is located in a district in which more than 60 percent of voters supported the ruling party in the

[1] The Ghana and Kenya DHS surveys are from 2003; the Malawi survey is from 2004; the Zambia survey is from 2007.

previous election) and ethnic group membership (defined by the ethnicity of the female household head).[2] We find that, both with respect to partisanship and ethnicity, the answer to the question "who is favored?" varies with the outcome being studied. We also demonstrate that determinations of political favoritism are highly sensitive to the modeling specifications one employs in one's analysis.

## Data

DHS data are highly advantageous for testing whether patterns of political or ethnic favoritism vary across public goods outcomes.[3] First, as noted, DHS surveys collect information about multiple outcomes in each survey household, which makes comparisons across outcomes straightforward. Second, the fact that DHS surveys have been implemented in standardized fashion in multiple countries makes it possible to undertake our analysis in several settings, which increases our confidence in the generalizability of our findings. Third, since the mid-1990s, DHS enumerators have recorded the geographical coordinates of enumerated households. This makes it possible for us to assign households to electoral constituencies (critical for our investigation of partisan targeting), calculate each household's distance from the capital city (which we use as an important control in our analyses), and, in combination with maps of lights visible from space at night, determine whether households live in lit clusters (which we use to supplement the electrification measure reported in the household surveys). Fourth, DHS surveys collect information about numerous characteristics of each household (urban/rural location, assets, housing material) and household head (age, literacy, religion, and ethnicity). This information makes it possible for us to condition our estimates of favoritism on need, which, in turn, allows us to disentangle the political component of the targeting we observe from the social planning component. The intuition here is that an apolitical social planner would allocate public goods to those households that need them most. So, once we have controlled for need, we would expect to observe no favoritism along partisan or ethnic lines. Whatever differences we do observe across partisan or ethnic groups in the control-laden specifications can thus be taken to be evidence of political favoritism. Finally, the DHS data has the advantage of permitting a household-level analysis, which allows us to avoid relying on ecological inferences about the relationship between partisanship/ethnicity and public goods provision—a problem that plagues many studies in this literature.

Despite these advantages, the data we employ do have several drawbacks. The most important of these is that the DHS surveys do a much better job of measuring stocks of public goods than flows. This is a problem since distributive politics is about the channeling of resources to constituencies for political ends—that is, about flows rather than stocks. While these resources may build up over time within targeted communities, the predictions of the literature are much more clearly about transfers of goods than about their accumulation. Furthermore, stocks of some durable goods such as infrastructure tend to persist over time, which makes it difficult to disentangle recent preferential treatment from past favoritism.

[2] Because the designers of the DHS are interested in fertility issues, they sample more women than men. We therefore employ the female surveys, which have a larger sample size. Hence we identify household ethnicity from the female, rather than male, household head.
[3] Franck and Rainer (2009) also take advantage of DHS data to identify patterns of ethnic favoritism in Africa.

Of the five outcomes we analyze, our measures of educational attainment and access to improved water sources are most vulnerable to this criticism (although we take some solace in the fact that our educational attainment results are similar when we break the respondents into age cohorts and conduct the analyses separately for each).[4] Our measure of vaccinations, which records the number of major childhood vaccinations that the household's youngest member has received is less vulnerable to this problem, as it picks up a valuable good received in the recent past, and is thus the most "flow-like." Our measures of access to electricity (based both on the DHS survey and on our separate analysis of the night lights data) fall somewhere in between, as the infrastructure required to deliver electricity to a household constitutes a major, durable, fixed investment (and is thus best thought of as a stock) but whether or not electricity is actually flowing at a given moment may be a product of a much more proximate political decision (which makes it—literally, in fact—a flow) (Min 2009).

The five outcomes we study also differ in the extent to which they can be seen unambiguously as the result of allocation decisions made by the government. Whereas electricity delivery and education are clearly products of government action, access to improved water and childhood vaccinations are often products of interventions by donors or NGOs. To the extent that we want to draw conclusions about political favoritism from the distribution patterns we observe with respect to these outcomes, the question of who makes the decisions about how they are to be distributed clearly matters. If donors or NGOs are deciding who gets vaccinated or where new boreholes are to be dug, then we would be wrong to interpret the receipt of vaccinations or boreholes by members of a particular ethnic group or by people living in an area where a particular political party is strong as evidence for distributive politics. However, precisely because things like vaccinations and water improvements are highly valued, governments usually attempt to direct their distribution—often by telling donors and NGOs where they can and cannot work—to their favored constituents. We interpret our findings with respect to vaccinations and access to improved water sources in this light.

The coding of our five dependent variables is as follows (see the tables in the Appendix for descriptive statistics by country and ethnic group):

1. *Household electricity*: 0 if no household electricity, 1 if household electricity.

2. *Lit Cluster*: 0 if person lives in a sampling cluster with no light output, 1 if person lives in a cluster with any light output.

3. *Water Source*: 1 if the water source is natural, such as from rainwater, a lake, stream, or pond; 2 if the source is an unprotected borehole or well; 3 if the source is a protected borehole or well; and 4 if he respondent's water source piped or better (bottled and so on).

[4] Respondent's ages run from 15 to 50. Separating the respondents into 5-year age cohorts (15-20, 20-25, 25-20, and so on) and testing for evidence of favoritism in each cohort does not result in substantively different findings. In future versions of this research we will more clearly link cohorts in each country to regimes that would have been in power during their school-age years.

4. *Education Level*: 1 if no education; 2 if only primary; 3 if secondary; 4 if higher.

5. *Number of Vaccinations*: measures the number of vaccinations that the youngest child in the household has received (0-4), where the potential vaccinations are BCG, DPT, polio, and measles.

**Empirical Strategy**

To test for patterns of political and ethnic favoritism with respect to each of our five outcomes, we run separate logistic (or ordered logistic) regressions for each outcome in which the main explanatory variable is a dummy variable indicating whether the household in question is located in a core district of the ruling party (which we define as a district in which more than 60 percent of voters supported the ruling party in the previous election) or is a member of one of each country's four major ethnic groups.[5] We code membership in an ethnic group in two ways: in terms of the ethnic group membership of the household head, and in terms of the predominant ethnicity of the sampling cluster in which the household is located.[6] The interpretation of the coefficient estimate on this dummy variable is that it captures the extent to which members of the partisan or ethnic group in question differ from members of other partisan or ethnic groups in the country. Significant positive coefficients thus connote favoritism, whereas significant negative coefficients connote that they category is disfavored vis-à-vis others.

The regression models include controls for the household head's age, literacy[7] and religion; whether the household is located in a rural area; as well as measures of household assets, the material used to make the household's floor, and the distance from the country's capital of the sampling cluster in which the household is located. As noted, our rationale for including these controls is to generate estimates of favoritism that control for need (and, through the "distance from the capital" covariate, also for the feasibility of providing the public good in question). Each of the models presented in this section also include controls for the region or province in which the respondent lives. The rationale here is to control for the possible spillover effects resulting from favoritism directed at one group being enjoyed by members of another group that happens to live nearby. We return in the last section of this Note to the sensitivity of our findings to alterations in the controls we employ.

Thus, if we are interested in learning whether the Kalenjin in Kenya are favored in their access to household electricity we run the following regression:

$$\Pr(Electricity = 1) = \log it^{-1}(\beta_0 + \beta_1 Kalenjin + \beta X)$$

[5] In Ghana, these are the Akan, Ewe, Ga and Mole-Dagbani. In Kenya, they are the Kalenjin, Kikuyu, Luhya and Luo. In Malawi, they are the Chewa, Ngoni, Tumbuka and Yao. In Zambia, they are the Bemba, Lozi, Nyanja, and Tonga.
[6] This latter specification drops households that do not reside in homogeneous sampling clusters, which we define as clusters in which more than 70% of household heads are members of a single ethnic group.
[7] We omit the literacy control in the educational attainment regressions.

where $X$ is a vector that includes the control variables described above. Similarly, if we are interested in whether individual's living in core constituencies of the UDF in Malawi are more likely to live in a lit cluster, we run the following regression:

$$\Pr(LitCluster = 1) = \log it^{-1}(\beta_0 + \beta_1 UDF + \beta X)$$

Because our strategy involves running 100 separate regressions (5 groups per country * 5 outcomes per group * 4 countries), and because we are only interested in inferences about the degree to which each group is favored relative to others living in each country, we only report here the coefficients and 95 percent confidence intervals for the coefficients of direct interest. The 95 percent confidence intervals from the models that use dummy variables for the political party core constituencies are computed using clustered robust standard errors.

**Results**

Figures 1-5 present the results from these analyses. Each panel in each figure presents the inferences that one would make about ethnic or political favoritism for each of the outcomes that we study. The points are logistic or ordered logistic regression coefficients and the lines are the 95 percent confidence intervals.

Figure 1 focuses on partisan favoritism. For each country, we can see whether households located in areas of core support for the ruling party are favored with respect to each of the five outcomes that we study. The findings are broadly consistent with our claim that the answer to the "are ruling party supporters favored?" question depends on the outcome one happens to be investigating. In Kenya, for example, the children of households located in districts where over 60 percent of voters had previously voted for KANU received a greater number of early childhood vaccinations than other Kenyans.[8] Yet these same households are less likely to have access to an improved water source or to live in a lit cluster. Thus, whereas a study of political favoritism in Kenya that happened to seize upon data on childhood vaccinations as its indicator of favoritism would conclude that KANU did, in fact, direct valuable patronage goods toward its core supporters, a study that focused instead on the distribution of electrification or access to improved water would reach the opposite conclusion.

Similarly, in Malawi, individuals living in core UDF districts are no more likely than other Malawians to have household electricity or to live in a lit cluster, yet they are, all else equal, more likely to have access to improved water. Thus, if one focused one's study of Malawian distributive politics exclusively on electrification one would conclude that UDF supporters are not favored. But if one focused instead on access to clean water one would conclude that UDF supporters are favored disproportionately, even controlling for their geographic location and level of need.

[8] Note that the controls included in the model permit us to rule out the possibility that the children of KANU supporters received more vaccinations because they were more in need of them—at least insofar as measures of household assets, literacy, rural location, etc are correlated with the need for vaccinations.

Our results on ethnic favoritism provide even stronger evidence for the proposition that inferences about preferential treatment of particular groups depend critically on the outcome that one chooses to study. Figures 2-5 present these results by country for four selected ethnic groups. Figures 2b, 3b, 4b, and 5b present results from analyses on a subset of the data that restricts the sample to respondents living in sampling clusters in which at least 70 percent of those surveyed are members of the same ethnic group. As government's often favor (or disfavor) ethnic groups by targeting the areas where they are known to live in high concentrations, we conduct these analyses as a check on our household-level results from the full dataset.

Figures 2a and 2b present results from Ghana. Of the four countries we examine, the findings from Ghana provide the weakest evidence in favor of our argument. Nevertheless they still challenge the notion that inferences about favoritism can be confidently drawn from the study of only one or a few outcomes. For example, if one were to study the Ewe, the ethnic group of former-president Jerry Rawlings and the group often assumed to have benefited most during his two decades in power (1981-2000), one would conclude that they have been substantially disadvantaged in their access to clean water relative to individuals of similar need. A study of vaccinations or educational attainment, however, would lead to the conclusion that the Ewe have been neither advantaged nor disadvantaged relative to other Ghanaians. A study of electrification infrastructure in Ghana would conclude that the Ga, the ethnic group that has historically lived in the areas in and around Accra, are quite substantially favored, even once we have accounted for their proximity to the capital. Studies of access to education or improved water, however, would conclude that the Ga have been disfavored.

Results from Kenya (Figures 3a and 3b) strongly support the notion that inferences about distributive favoritism vary strongly with the good that one happens to study. Kalenjin households, for instance, appear to be favored with respect to their access to vaccinations, disadvantaged with respect to their access to clean water, and comparable to other similar Kenyans with respect to their access to electricity. For their part, Luo households appear to be favored with respect to their access to improved water but either disadvantaged or similar to other Kenyans with respect to their access to the other outcomes that we study. A study of access to improved water sources in Kenya would thus find the Kikuyu and Luo to be advantaged, the Kalenjin to be disadvantaged, and the Luhya to be neither advantaged nor disadvantaged. A political story about distributive favoritism could then be readily developed that would be consistent with this pattern. Yet a study of access to vaccinations would find the Kalenjin and Luhya to be favored, the Kikuyu to be disfavored, and the Luo to be neither disfavored nor favored, a pattern that would require a quite different political explanation.

Similar findings emerge from our analysis of distributive patterns in Malawi (Figures 4a and 4b) and Zambia (Figures 5a and 5b). Though in a few instances ethnic groups appear favored in their access to almost all goods—the Ngoni in Malawi, for instance, have positive coefficients in four of the five outcomes—the inferences one would draw about the degree to which ethnic groups are favored generally vary with the goods upon which one focuses. The Chewa in Malawi, for example, appear disfavored with respect to their access to clean water but are comparable to other Malawians in their access to vaccinations and electricity. Similarly, in Zambia, one would conclude that Bemba speakers are disfavored in their access to clean water

and vaccinations but neither favored nor disfavored in their access to education and electricity. The clear conclusion is that the answer to the "who is favored?" question depends on the particular outcome on which one happens to focus.

## The Importance of Model Specification

In the previous section, we justified our control-laden specifications by pointing out that, because our interest lies in the political rationale behind allocation decisions, it was sensible to control, as much as possible, for need and for the possibility of unintentional spillovers from favored groups to unfavored groups who happen to live in close proximity. In this section we return to this issue to explore in greater detail the implications of this decision and, in doing so, demonstrate the sensitivity of one's findings to the reference group that, when one selects a particular model specification, one (implicitly) imposes on one's estimates of favoritism.

The answer to the "who is favored?" question is a relative judgment. It implies a second question—"favored relative to whom?"—the answer to which has important implications for the appropriate specification one should use in one's analysis. If, for example, we are interested in absolute favoritism (or well-being) then we would want to study unconditional differences between groups. The potential drawback is, of course, that such a raw measure risks confusing historical disadvantage with current discrimination. If we are interested in comparing access to goods relative to other people with similar needs, then we would want to attempt to control for these needs in the analysis. And if we are interested in ruling out "inadvertent" benefits—deriving from the non-excludability of what are effectively local public goods—or in comparing individuals more directly to others in the places where they live then we would also want to integrate geographic controls (in our case, regional or provincial) into the analysis.

Figure 6 demonstrates that these specification decisions are not trivial, and that the inferences one draws about ethnic and political favoritism can be very sensitive to the reference group imposed by one's analysis, and the set of controls that one therefore includes in one's regressions. The figure presents coefficient estimates and 95 percent confidence intervals from three different model specifications: one including only the political or ethnic group dummy (to capture unconditional differences); one including controls to capture differences in need across individuals; and one including the geographic controls as well. Even a cursory look at the Figure makes it clear that inferences about distributive favoritism can depend not only upon the outcome that one happens to study but also upon the reference group that one imposes in the statistical analysis.

Consider, for example, households located in the core constituencies of the National Democratic Congress (NDC) in Ghana (ostensibly the ruling party from 1981-2000). When one compares their access to improved water to that of other Ghanaians (Figure 6, row 3; the "no controls" point estimate—black circle), these individuals appear quite disadvantaged. Yet when we account for the fact that NDC supporters are generally poorer and compare them to individuals of similar need (the "controls" point estimate—blue triangle), they appear to have been favored in the provision of clean water. Interestingly, however, when compared to individuals of similar need in the same region, they do not appear favored ("controls and regional

dummies" point estimate—red X). This result likely emerges from the nature of regional partisan divisions in Ghana. There are broadly three kinds of regions in Ghana: those that are (nearly) homogeneously NDC supporting (in Volta and in the north); those that are (nearly) homogeneously NPP supporting (in Asante and in Brong Ahafo); and those that are mixed. The former provide little or no leverage in estimating intra-regional differences (there is little or no variation in partisanship) and so the entire estimate is being driven by differences between individuals in those regions where the parties both compete successfully in some districts. These are also the regions where the differences between NPP and NDC constituencies are likely the least pronounced and where we would expect the least politically motivated targeting toward core constituencies (compared to the constituencies in the homogenous regions—especially Volta, an NDC stronghold, and Ashante, an NPP stronghold—these areas are likely perceived as less "core" by each party).

A similar point can be made if one looks at partisan favoritism with respect to access to vaccinations in Ghana (Figure 6, row 9). Children in NPP core districts appear disadvantaged with respect to the number of vaccinations they receive from government ("no controls"—black circle). Yet when one accounts for the fact that NPP supporters tend to be wealthier that NDC supporters ("controls"—blue triangle), children in NPP districts appear favored. Importantly, the NPP took power two years before the DHS survey and the vaccination measure refers to the number of vaccinations received by the household's youngest child. This suggests that generally the period in which such vaccinations probably were received was the period after the NPP took power (as most vaccinations are received in a child's first year).

The "Northwesterners" in Zambia—that is, members of the Luvale, Kaonde, Lunda, and other ethnic groups living in the northwestern part of the country—provide similar evidence that inferences depend upon whom the reference group is that one selects in the analysis (Figure 6, row 1). In comparison to other Zambians, Northwesterners are substantially disadvantaged in their access to electrical infrastructure ("no controls"—black circle). But compared to Zambians of similar need ("controls"—blue triangle), they appear neither advantaged nor disadvantaged and, interestingly, compared to others in the regions in which they live (predominantly in the Northwestern Province), they appear substantially favored in their access to electrical infrastructure ("controls and regional dummies"—red X). Thus while Northwesterners appear disfavored in comparison to most Zambians, they appear favored relative to others in the areas in which they live. One explanation for this pattern may be that Northwesterers are the "titular" northwest people, and are more likely to be in the towns and larger communities within the northwest (as those towns/larger communities were built up around the chieftaincies of these groups). The result is a mechanical correlation between membership in these titular groups, relative urbanization, and thus the probability of living in a lit cluster.

The pattern with the Mole-Dagbani groups in northern Ghana is similar: they are the titular northerners and, within the northern provinces in which they predominate, are more likely to be concentrated in larger communities and towns. Thus, although the Mole Dagbani appear disfavored with respect to their access household electricity relative to other Ghanaians (Figure 6, row 2; "no controls"—black circle), Mole-Dagbani appear favored relative to other inhabitants of the country's northern provinces ("controls and regional dummies"—red X).

Figue 6 presents only an illustrative sample of cases where the conclusion one would reach about group favoritism happens to be particularly sensitive to the choice of specification. However, the clear implication is that researchers must carefully consider the comparison group when they evaluate patterns of political and ethnic favoritism in the provision of public goods. The answer to the question of "who is favored?" often depends critically on these considerations—as well as on the outcome that one chooses to study, as we demonstrated earlier.

## Conclusion

The empirical distributive politics literature is growing rapidly. As data become more readily available in the developing world, the literature is continually expanding from the original studies that predominantly focused on North America and Western Europe to such diverse locations as India, Guinea, and Mexico. Many of these studies are valuable contributions to our understanding of government behavior and distributive patterns in specific domains. Often, these specific areas of distribution and targeting will be profoundly important and worthy of focus in their own right, as is the case with such outcomes as infant and child mortality.

Yet if the goal of the research is to make general statements about governments' distributive strategies or propensities to favor certain political constituencies or ethnic groups, a narrow focus on one or a few goods can be problematic. Our findings from four African countries—all places where accusations of ethnic favoritism have been leveled at governments—suggest that extrapolating from inferences drawn from an analysis of one or a few sectors to conclusions about government favoritism *per se* may be misleading. Our results suggest, therefore, the need to examine a complete (or at the very least *more* complete) portfolio of government activities before drawing general inferences about who is favored by the country's distributive politics.

A second implication of our findings is that inferences about political and ethnic targeting are highly sensitive to the individuals or groups against which one (often) implicitly compares the group of interest. Put simply, the reference group in the analysis matters in ways that can completely flip the direction of one's inferences. Is a group disfavored if it has less access to a valued good than other groups but nonetheless greater access than groups that are geographically proximate? Is a group favored if it enjoys greater access to a particular good than the rest of the country but less access than groups with similar levels of need? These are the types questions that we faced in the course of the research, and they suggest that answering the question of who is favored is more subtle than simply measuring who has more. Our results thus point to the need for researchers to be explicit about what they mean by favoritism and then to make certain that the suite of control variables they include in their statistical models reflect the definitions they have decided upon.

Though our argument is simple, we believe that it poses a challenge to an important and growing literature in political science and economics. Understanding the logic of government distributive decisions is consequential, both for scholars interested in how electoral and identity politics shape government decision-making and because of the welfare implications of public

goods distribution. Through this research we hope to push the literature forward toward a more complete understanding of the political economy of public goods distribution.

## Table 1: Selected Recent Empirical Investigations of Distributive Politics

| Paper | Country | Outcome/Good Studied | Level of Analysis (N) |
|---|---|---|---|
| Alesina et al. (1999) | Italy | Public employment | |
| Arulampalam et al. (2009) | India | Federal grants to state governments | |
| Banful (2008) | Ghana | Allocations to local governments | Districts (110) |
| Barkan and Chege (1989) | Kenya | Roads spending<br>Health Spending<br>Rural Development Fund | Provinces (7) |
| Burgess et al. (2010) | Kenya | Roads | District |
| Calvo and Murillo (2004) | Argentina | Province expenditures financed by federal government | |
| Case (2001) | Albania | Block grants from federal to local government | |
| Castellsa and Sole-Ole (2005) | Spain | Infrastructure expenditures | |
| Crampton (2004) | Canada | Job grants | |
| Dahlberg and Johansson (2002) | Sweden | Temporary intergovernmental grant program | |
| de la Fuent and Vives (1995) | Spain | Infrastructure and education expenditures | |
| Denemark (2000) | Australia | Constituency level grants | |
| Diaz-Cayeros et al. (2008) | Mexico | Cash transfers, Water, Electricity | |
| Drazena and Eslava (2005) | Colombia | Municipal Budgets | |
| Frank and Rainer (2009) | Multiple African | Infant mortality and Education | Individual (over 300,000) |
| Gershon (1084) | Israel | Federal grants to local governments | |
| Golden and Picci (2008) | Italy | Infrastructure expenditures | |
| Gonzalez (2002) | Mexico | Cash transfers and infrastructure spending | |
| Govinda and Singh (2001) | India | Federal grants to state governments | |
| Horiuchi and Lee (2008) | South Korea | Pork barrel allocations | |
| Hugh and John (1999) | UK | Grants to local governments | |
| Hyeok (2005) | South Korea | Subsidies | |
| Johansson (2003) | Sweden | Intergovernmental grants to municipalities | |
| Kasara (2007) | Multiple African | Agricultural taxes | Ethnic groups (50) |
| Kudamatsu (2007) | Guinea | Infant mortality | Individual (13,837) |
| Miguel and Zaidi (2003) | Ghana | Education spending | Schools (192) |
| Milan and Sircar (2009) | India | Schools | |
| Milligan and Smart (2001) | Canada | Regional grants | |
| Porto and Sanguinetti (2003) | Argentina | Federal grants to local governments | |
| Posner and Simon (2002) | Zambia | Poverty | District (39) |
| Remmer (2003) | Argentina | Provincial level spending | |
| Sapienza (2004) | Italy | Government-owned bank interest rates | |
| Schady (2000) | Peru | Social Fund expenditures | |
| Sole-Olle et al. (2008) | Spain | Intergovernmental transfers | |
| Tavits (2009) | Nordic Countries | Federal grants to local governments | |
| Treisman (1996) | Russia | Intergovernmental transfers | |
| Winters (2009) | Ecuador | Cash transfers | |
| Worthington and Dollery (1998) | Australia | Intergovernmental grants | |
| Zucco (2008) | Brazil | Cash transfers (Bolsa Familia) | |

**Figure 1**: How inferences about favoritism toward core constituencies (areas where over 60 percent of voters voted for the ruling party) are influenced by the outcome selected to study. Explanatory variable of focus is dummy variable indicating whether the majority of the voters in the district supported the incumbent. Plots present logit or ordered logit coefficient on dummy variable indicating whether individual lives in a district in which a majority voted for the ruling party at the time of the survey. Lines show 95 percent confidence intervals computed using robust clustered standard errors.

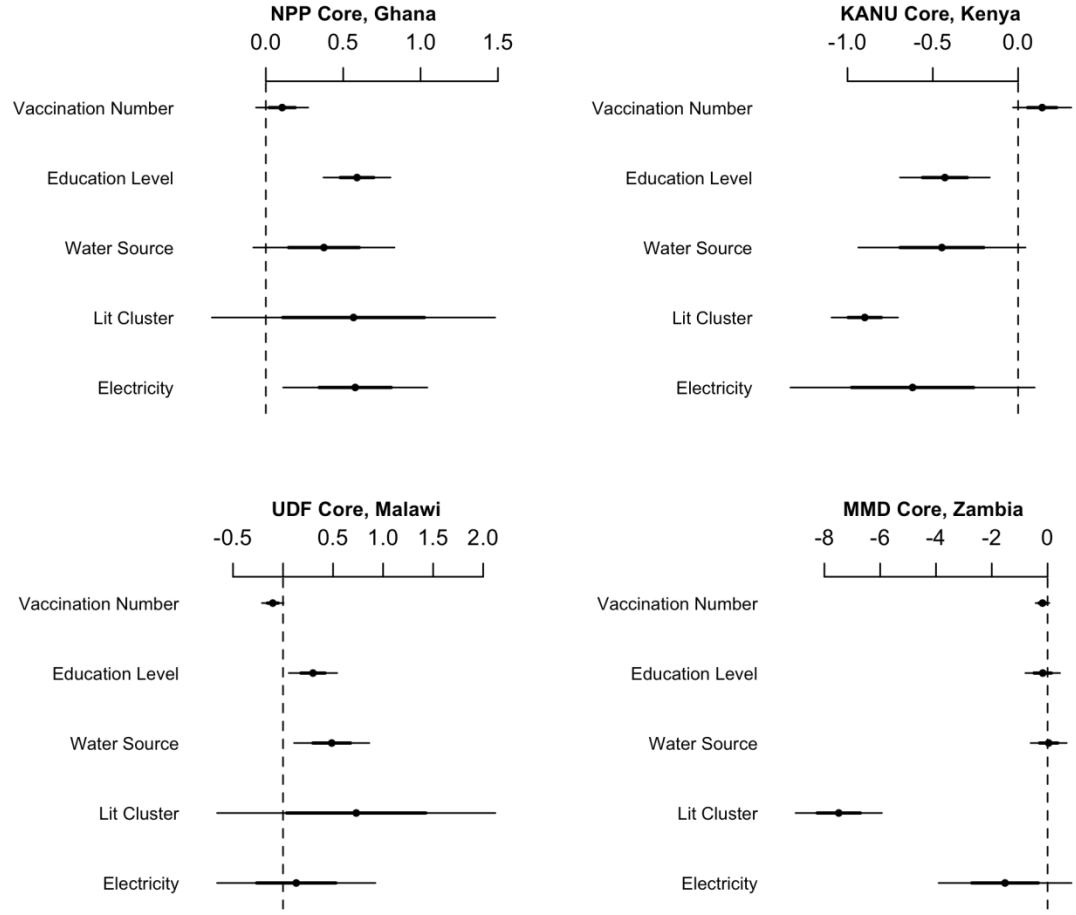


**Figure 2a**: Ghana: Inferences about ethnic favoritism by group. Plots present logit or ordered logit coefficient on dummy variable for ethnic group. The coefficient captures the difference between the group and the rest of the country. Lines show 95 percent confidence intervals.

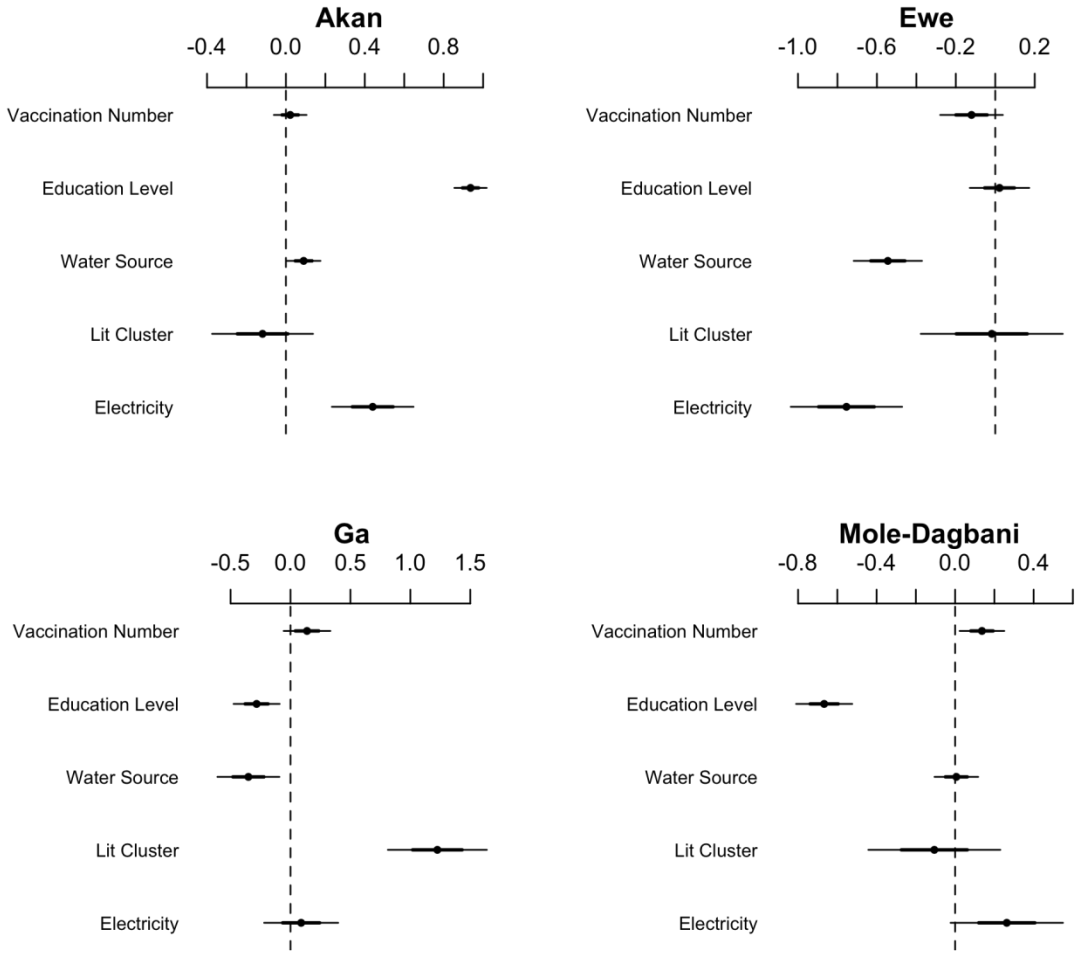


**Figure 2b**: Ghana, subset of ethnically homogenous clusters: Inferences about ethnic favoritism by group. Plots present logit or ordered logit coefficient on dummy variable for ethnic group. The coefficient captures the difference between the group and the rest of the country. Lines show 95 percent confidence intervals.

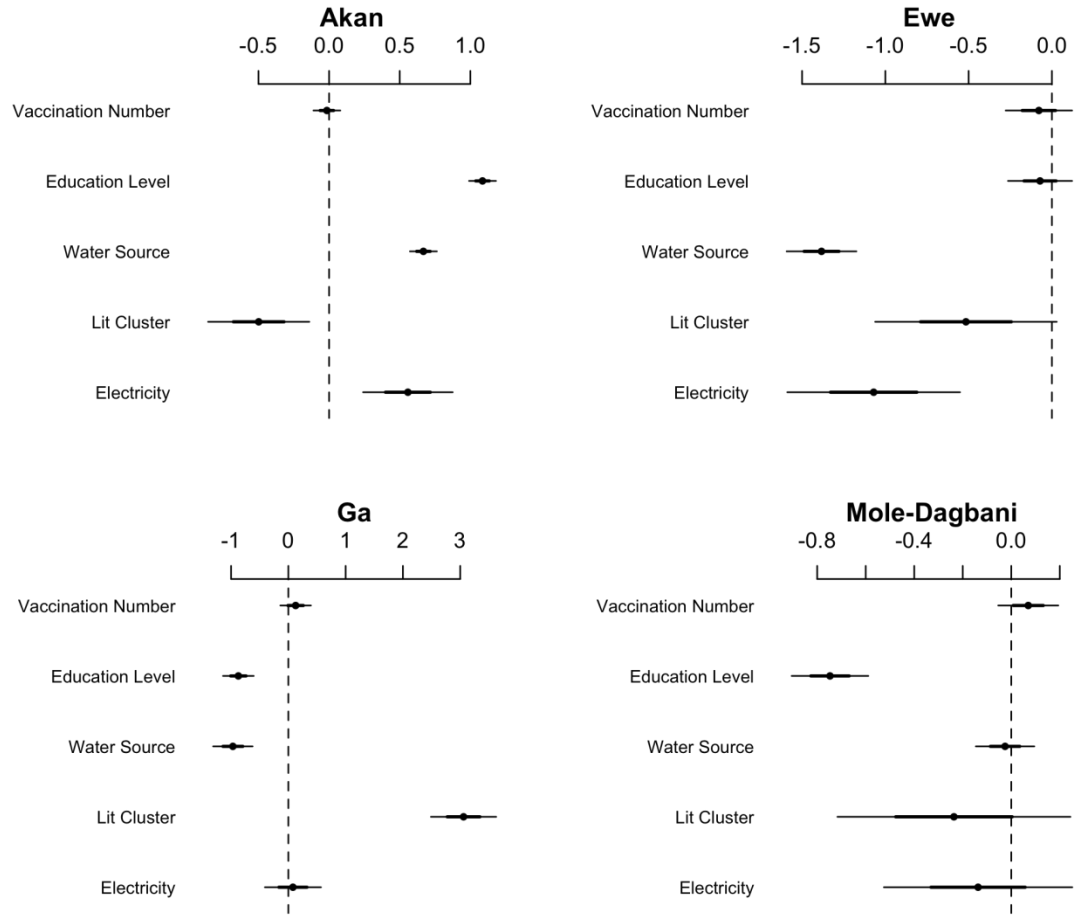


**Figure 3a**: Kenya: Inferences about ethnic favoritism by group. Plots present logit or ordered logit coefficient on dummy variable for ethnic group. The coefficient captures the difference between the group and the rest of the country. Lines show 95 percent confidence intervals.

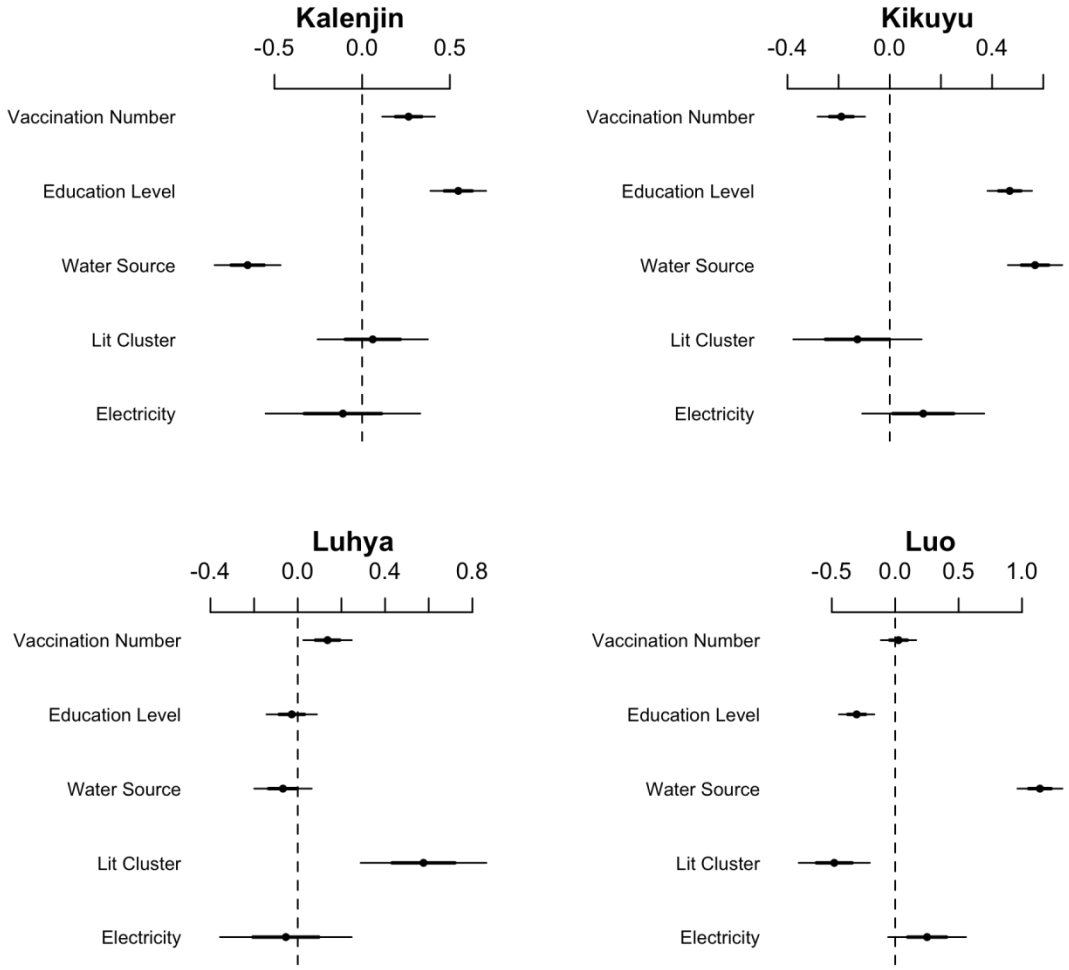


**Figure 3b**: Kenya, subset of ethnically homogenous clusters: Inferences about ethnic favoritism by group. Plots present logit or ordered logit coefficient on dummy variable for ethnic group. The coefficient captures the difference between the group and the rest of the country. Lines show 95 percent confidence intervals.

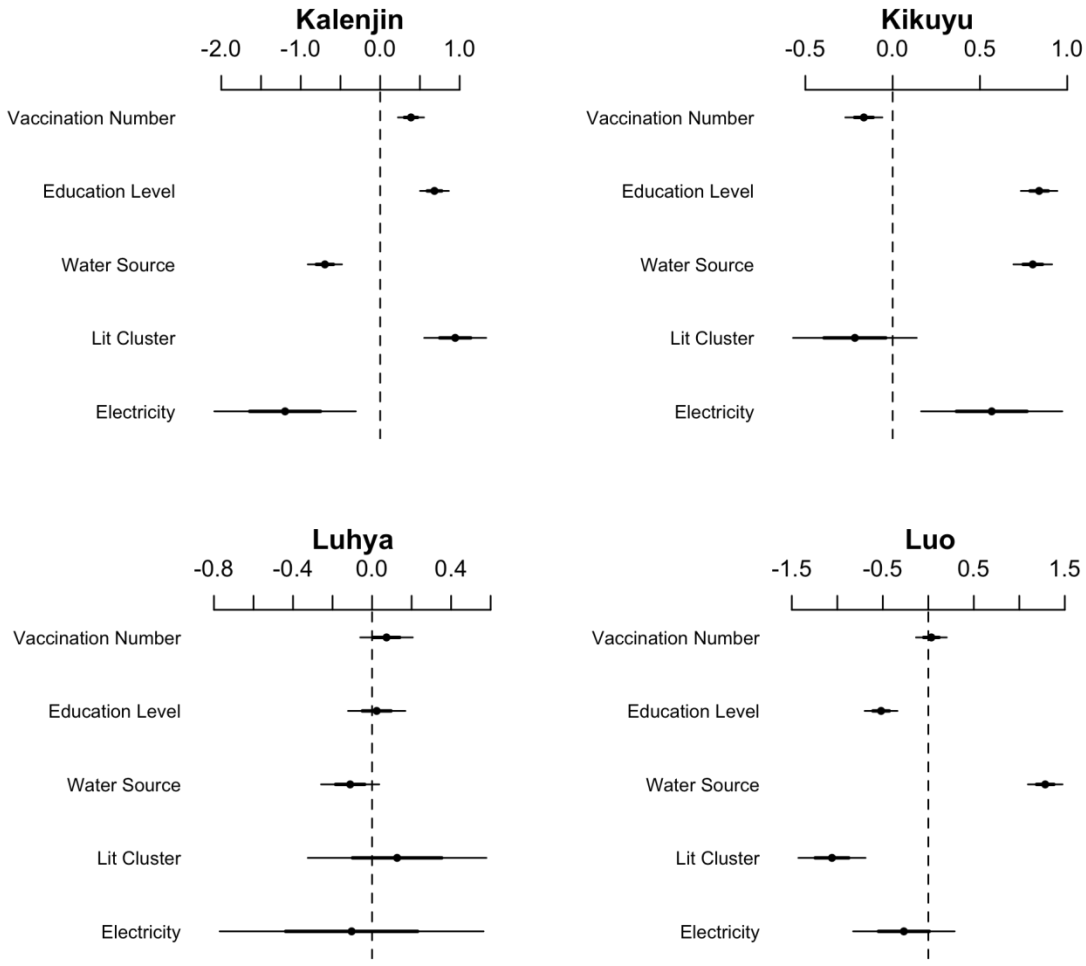


**Figure 4a**: Malawi: Inferences about ethnic favoritism by group. Plots present logit or ordered logit coefficient on dummy variable for ethnic group. The coefficient captures the difference between the group and the rest of the country. Lines show 95 percent confidence intervals.

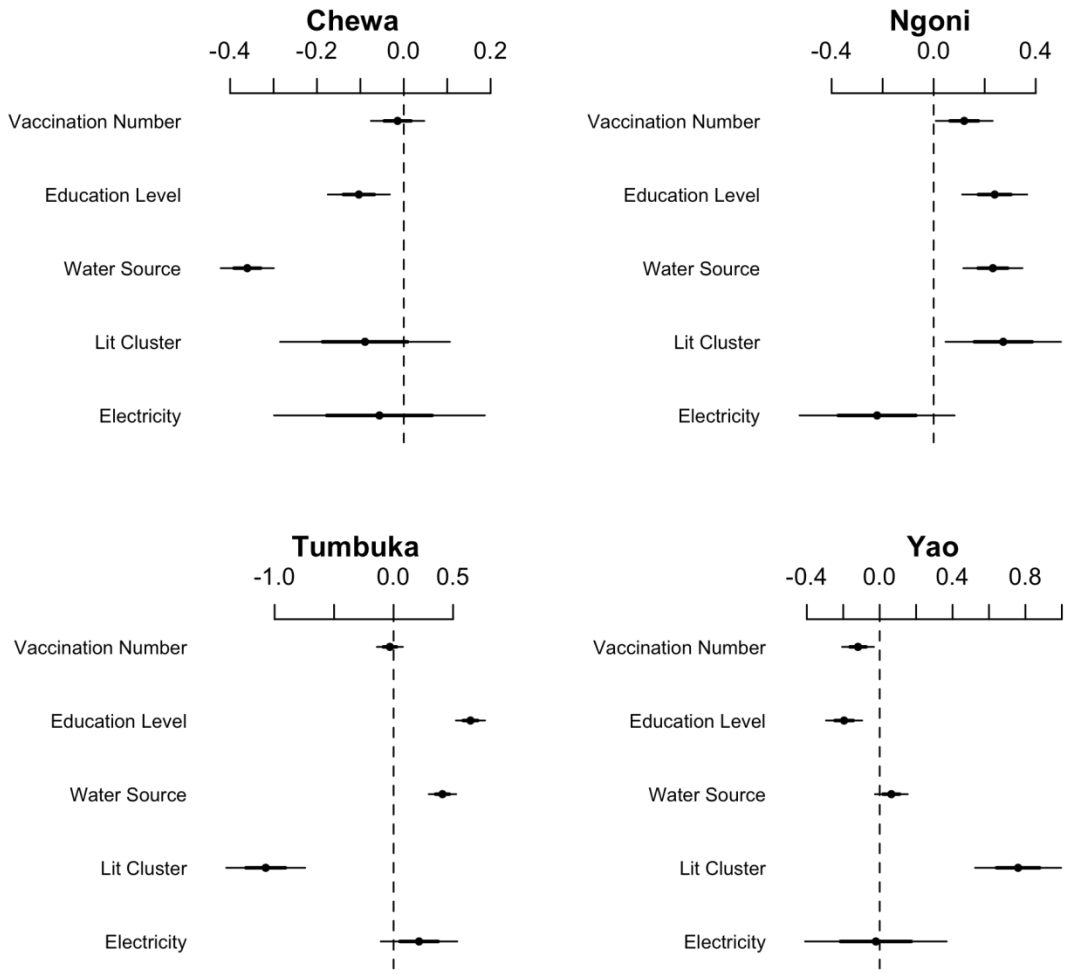


**Figure 4b**: Malawi, subset of ethnically homogenous clusters: Inferences about ethnic favoritism by group. Plots present logit or ordered logit coefficient on dummy variable for ethnic group. The coefficient captures the difference between the group and the rest of the country. Lines show 95 percent confidence intervals.

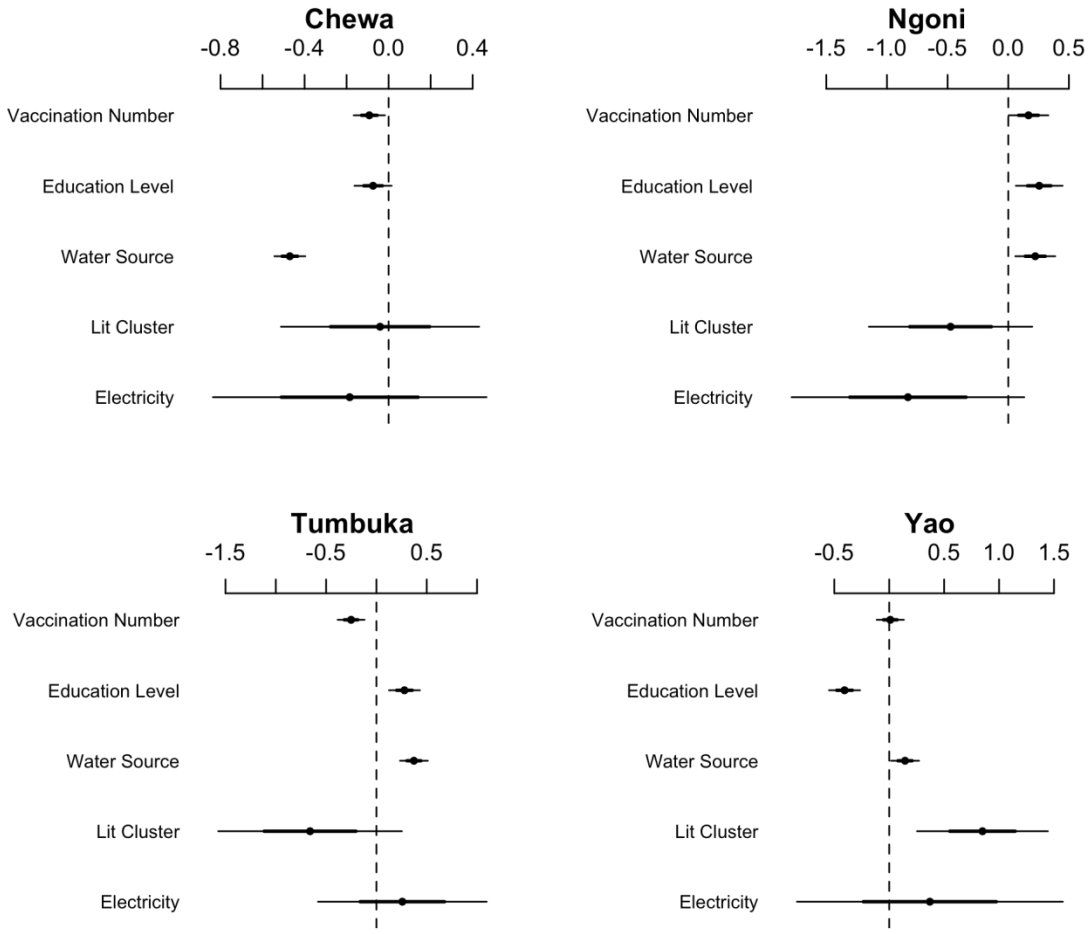


**Figure 5a**: Zambia: Inferences about ethnic favoritism by language group. Plots present logit or ordered logit coefficient on dummy variable for ethnic group. The coefficient captures the difference between the group and the rest of the country. Lines show 95 percent confidence intervals.

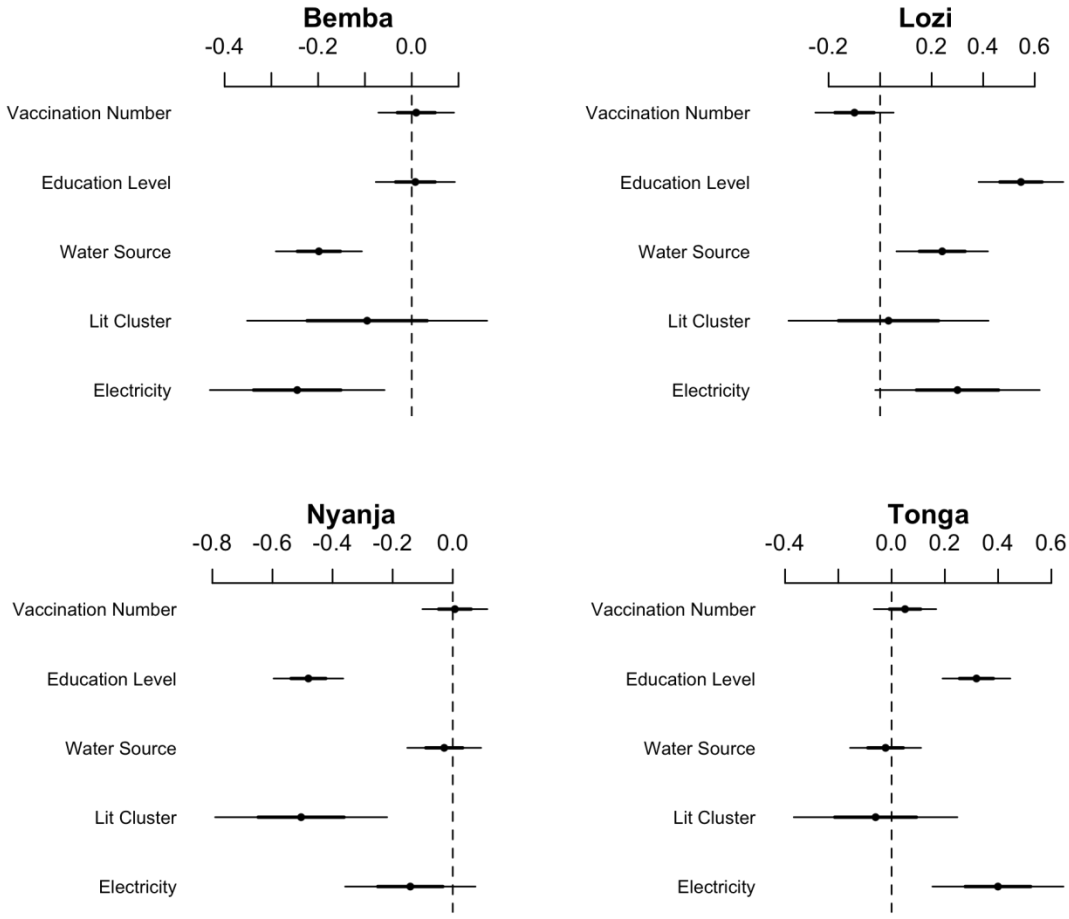


**Figure 5b**: Zambia, subset of ethnically homogenous clusters: Inferences about ethnic favoritism by language group. Plots present logit or ordered logit coefficient on dummy variable for ethnic group. The coefficient captures the difference between the group and the rest of the country. Lines show 95 percent confidence intervals.

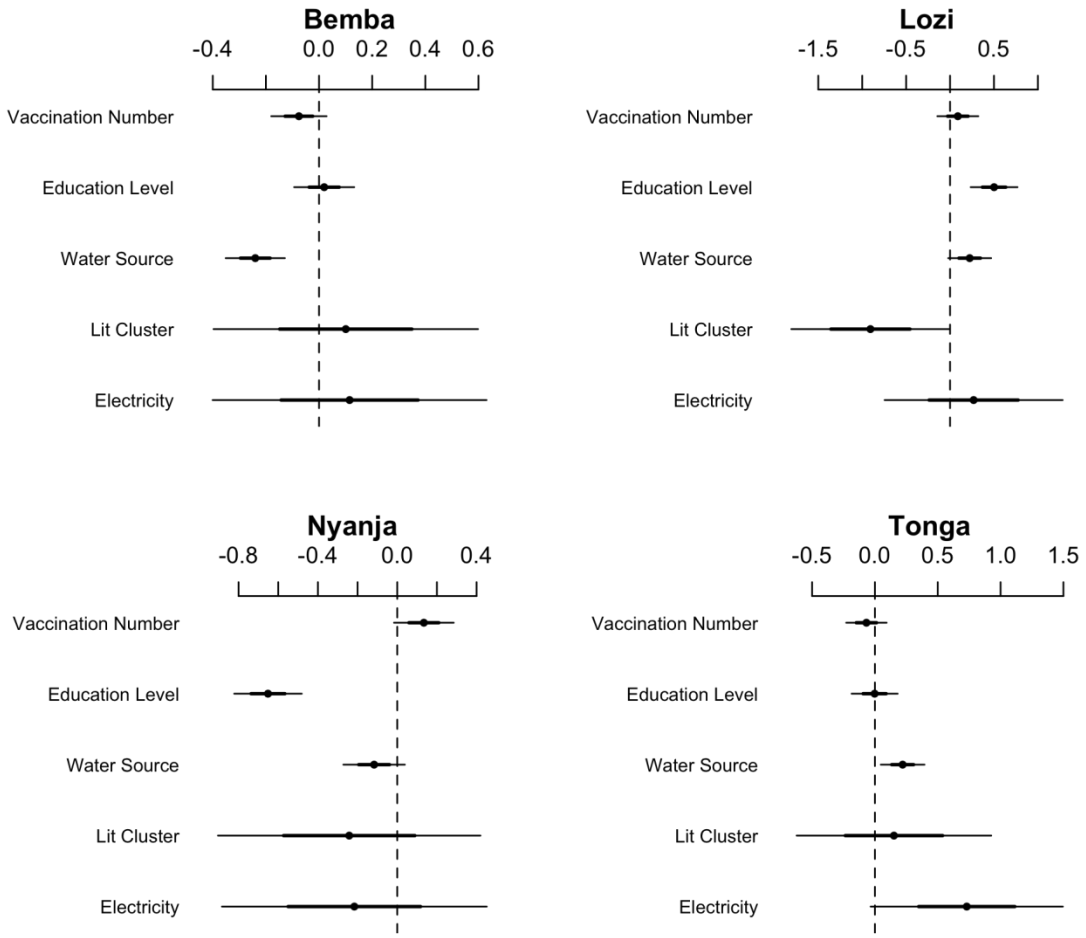
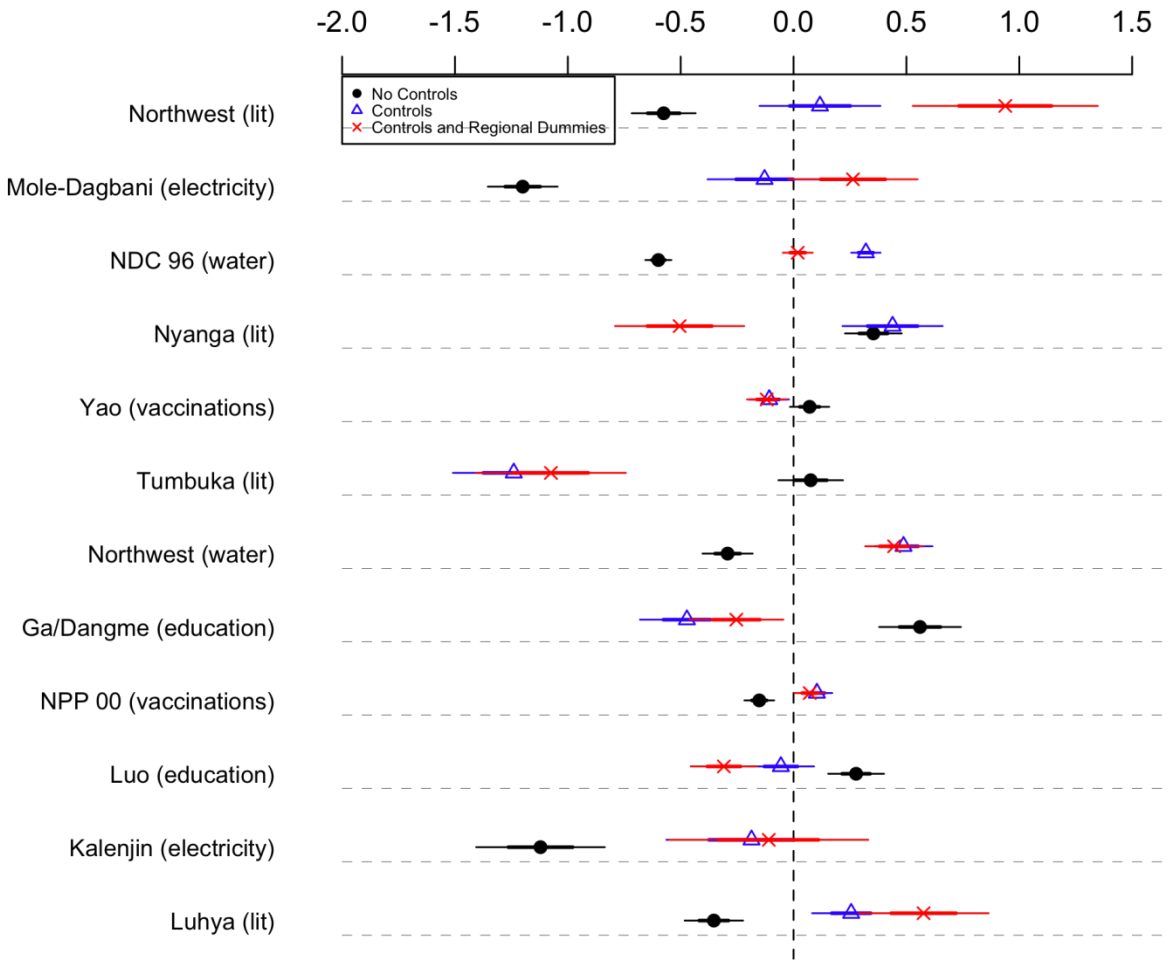


**Figure 6**: Inferences by Selected Groups and Goods. Inferences are Sensitive to Model Specification. Solid dot is inference from model without any controls. Triangle (blue) are for models with controls. Cross (red) is for models with controls and regional dummies.



**References**

Dahlberg, Matz and Eva Johannson. 2002. "On the Vote Purchasing Behavior of Incumbent Governments." *American Political Science Review* 96, 1: 27-40.

Diaz-Cayeros, Alberto, Beatriz Magaloni and Frederico Estevez. 2008. *Strategies of Vote-Buying: Poverty, Democracy, and Social Transfers in Mexico.* Unpublished book manuscript.

Franck, Raphael and Ilia Rainer. 2009. "Does the Leader's Ethnicity Matter? Ethnic Favoritism, Education and Health in Sub-Saharan Africa." Unpublished paper.

Golden, Miriam and Lucio Picci. 2008. "Pork-Barrel Politics in Postwar Italy, 1953-94." *American Journal of Political Science* 52, 2: 268-89.

Kasara, Kimuli. 2007. "Tax Me If You Can: Ethnic Geography, Democracy and the Taxation of Agriculture in Africa. *American Political Science Review* 101, 1: xx-xx.

Min, Brian. 2009. FILL THIS IN…

## Appendix

**Table A-1: Ghana, Variable Means by Ethnic Group**

| | Akan | Ga | Ewe | Mole-Dagbani |
|---|---|---|---|---|
| Proportion in Sample | 0.43 | 0.07 | 0.12 | 0.19 |
| Household Electricity | 0.57 | 0.56 | 0.37 | 0.22 |
| Lit Cluster | 0.62 | 0.76 | 0.53 | 0.19 |
| Water Source | 3.05 | 3.13 | 2.67 | 2.56 |
| Education Level | 2.53 | 2.41 | 2.37 | 1.49 |
| Vaccination Number | 1.39 | 1.39 | 1.27 | 1.78 |
| Rural | 0.49 | 0.47 | 0.59 | 0.75 |
| Literacy | 0.37 | 0.40 | 0.39 | 0.13 |
| Distance from Accra | 180.58 | 59.93 | 138.72 | 496.73 |
| Primary Region | Ashanti | G. Accra | Volta | Upper West |
| Primary Region Proportion | 0.28 | 0.59 | 0.47 | 0.35 |

**Table A-2: Kenya, Variable Means by Ethnic Group**

| | Kalenjin | Kikuyu | Luhya | Luo |
|---|---|---|---|---|
| Proportion in Sample | 0.08 | 0.24 | 0.15 | 0.10 |
| Household Electrification | 0.08 | 0.32 | 0.15 | 0.27 |
| Lit Cluster | 0.16 | 0.56 | 0.34 | 0.45 |
| Water Source | 1.78 | 2.81 | 2.22 | 2.75 |
| Education Level | 2.21 | 2.55 | 2.29 | 2.33 |
| Vaccination Number | 1.72 | 1.36 | 1.51 | 1.28 |
| Rural | 0.71 | 0.64 | 0.76 | 0.70 |
| Literacy | 0.69 | 0.86 | 0.74 | 0.79 |
| Distance from Nairobi | 232.84 | 77.02 | 259.60 | 221.79 |
| Primary Region | Rift Valley | Central | Western | Nyanza |
| Proportion in Primary Region | 0.81 | 0.57 | 0.64 | 0.58 |

**Table A-3: Malawi, Variable Means by Ethnic Group**

| | Chewa | Ngoni | Tumbuka | Yao |
|---|---|---|---|---|
| Proportion in Sample | 0.38 | 0.09 | 0.10 | 0.15 |
| Household Electrification | 0.06 | 0.09 | 0.15 | 0.05 |
| Lit Cluster | 0.16 | 0.27 | 0.24 | 0.26 |
| Water Source | 2.49 | 2.79 | 2.81 | 2.70 |
| Education Level | 1.89 | 2.02 | 2.22 | 1.71 |
| Vaccination Number | 1.87 | 1.96 | 1.83 | 1.98 |
| Rural | 0.88 | 0.81 | 0.79 | 0.89 |
| Literacy | 0.51 | 0.61 | 0.74 | 0.40 |
| Distance From Lilongwe | 106.79 | 161.66 | 260.37 | 193.74 |
| Primary Region | Central | Central | Northern | Southern |
| Proportion in Primary Region | .75 | .50 | .75 | .75 |

**Table A-4: Zambia, Variable Means by Ethnic Group**

| | Bemba | Lozi | Nyanja | Tonga |
|---|---|---|---|---|
| Proportion in Sample | 0.31 | 0.08 | 0.17 | 0.14 |
| Household Electrification | 0.24 | 0.27 | 0.26 | 0.26 |
| Lit Cluster | 0.47 | 0.40 | 0.52 | 0.41 |
| Water Source | 2.60 | 2.96 | 2.83 | 2.71 |
| Education Level | 2.34 | 2.43 | 2.24 | 2.40 |
| Vaccination Number | 1.37 | 1.46 | 1.43 | 1.55 |
| Rural | 0.50 | 0.53 | 0.52 | 0.67 |
| Literacy | 0.56 | 0.66 | 0.51 | 0.60 |
| Distance to Lusaka | 422.53 | 450.60 | 326.86 | 172.38 |
| Primary Region | Luapula | Western | Eastern | Southern |
| Primary Region Proportion | 0.29 | 0.69 | 0.54 | 0.56 |