Dear WGAPE participants,

We apologize for getting this draft to you so late in the game and hope you will have time to at least give it a quick read before we discuss it on Saturday. What we are circulating is a very preliminary first draft of a project we have been working on for some time. We are still working through the logic (and the corresponding appropriate operationalization) of the cross-constituency analysis, so we would be particularly interested in your feedback on that part of the paper. This said, any and all comments on the project would be most welcome!

Best wishes,

Dan and Andy

(Under What Conditions) Do Politicians Reward Their Supporters? Evidence From Kenya's Constituencies Development Fund*

J. Andrew Harris
NYU-Abu Dhabi

Daniel N. Posner
UCLA

7 January 2015

Preliminary Draft

Please Do Not Quote or Cite

*Prepared for the Winter 2015 WGAPE meeting, University of Washington, Seattle.

A broad literature examines how politicians allocate resources among their constituents with an eye toward re-election. In Africa, such processes are almost always viewed through the lens of ethnic patronage or clientelism, with the basic credo "It's our turn to eat" implying an allocative strategy akin to the core voter hypothesis: when in office, reward your co-ethnics and exclude your enemies (Wrong 2009). Empirical studies of this process, however, focus almost exclusively on the allocation decisions of national-level political actors to districts (e.g., Barkan and Chege 1989; Burgess et al 2014; Jablonski 2014). This paper contributes to the study of distributive politics in Africa by examining a particular policy innovation in Kenya, the Constituencies Development Fund (CDF), which permits us to study the targeting of political resources by more local actors—Members of Parliament (MPs)—at the within-district level.[1]

Since its inception in 2003, Kenya's CDF program has provided MPs with millions of dollars of funding to build development projects in their constituencies. Since MPs have complete discretion over where to locate the projects they fund with CDF monies, we can use the spatial allocation of such projects to test whether their placement is shaped by where their supporters are located. We explore those allocation decisions using a unique spatial dataset of more than 34,000 CDF projects initiated between 2003 and 2007. To explain the patterns we observe, we combine these data with voting returns from more than 14,000 polling stations in the 2002 parliamentary elections and highly disaggregated data on population density, the location of roads, poverty rates, and ethnic demographics. We then use spatial modeling techniques to model the placement of projects within each constituency as a function of these independent variables, generating estimates of the relationship between project location and political support in each of the 210 constituencies. Taken together, these data and methods put us in a unique position to learn about how politicians use the resources at their disposal to maximize the likelihood of re-election, and how the strategies they pursue vary with the local conditions they face.

We find evidence that MPs do in fact favor their supporters in the distribution of CDF projects. However, we find that MPs consider other factors as well, such as the number of people living in the area in question, the distance from a paved road, the poverty rate, and the proximity to their own village. More important, we find that the tendency for MPs to channel CDF projects to their supporters, while evident in general, varies from constituency to constituency. We find that this cross constituency variation is driven by the MP's margin of victory in the last election, the constituency's level of ethnic heterogeneity, and the extent to which the MP's supporters and non-supporters are spatially segregated from one another.

In its focus on district-level allocations, existing work on distributive politics in Africa implicitly assumes that all the constituents in a district all have equal access to the "pork" that is funneled to that district. This assumption ignores the fact that where a project is built often determines who can use it—for example, the benefit of a primary school or a borehole accrues primarily to constituents living near such infrastructure. So simply counting the number of schools or boreholes in a district may not tell us very much about

[1] This is important insofar as the leading theoretical treatments of distributive politics (Dixit and Londregan 1996; Cox and McCubbins 1986) focus on within-district rather than cross-district allocations.

"who gets what." By geo-referencing every CDF project, our approach allows us to directly examine the spatial allocation of development resources. It also allows us to avoid the problem, noted by Cox (2006), of conflating the individual MP's re-election concerns with his party's concerns regarding legislative outcomes and other national-level objectives. The CDF in the Kenyan context avoids this problem since CDF allocations by individual legislators are set independently from party-level concerns.

Our approach and results also allow us to further specify the theoretical bases for the core voter model by showing how the geo-spatial distribution of political support conditions a politician's ability to target rewards to supporters. Theories of distributive politics are strangely silent on the question of feasibility: can a politician actually target supporters with the given distributive technology at his disposal (i.e., public goods)? Whether or not he can may depend on whether supporters and opponents live in close proximity to one another. Our fine grained, geo-coded data permit us to study this issue in a novel way.

The Constituencies Development Fund (CDF) Program in Kenya

Created under the Constituencies Development Fund Act of 2003 (GoK 2003), the Kenya Constituencies Development Fund is provided an annual allocation of not less than 2.5 percent of all ordinary government revenue. These funds are then distributed among the country's 210 electoral constituencies according to a formula by which 75 percent are divided equally and the remaining 25 percent are allocated based on each constituency's contribution to national poverty. Once the funds arrive in each constituency, they are disbursed by the constituency-level CDF Committee, which is controlled by the MP.[2] Citizens and organized groups may (and do) apply for projects, but the Committee (in practice, the MP) determines which projects are funded and where they are located.

CDF transfers represent a considerable infusion of funds for local development. Total CDF allocations between 2003 (when the program was introduced) and 2007 totaled nearly $370 million (see Table 1). This amounts to an average of $351,900 per constituency per year. These allocations funded a total of 34,139 projects, or an average of about 163 projects per constituency over this five-year period.[3]

(Table 1 Here)

[2] CDF funds are widely regarded as *pesa ya Mheshimiwa*, Swahili for "the MP's money" (Institute of Economic Affairs 2006; see also Awiti 2008 and National Anti-Corruption Campaign Steering Committee 2008). The MP's ability to control the CDF Committee stems from his statutory role as its chairman, his ability to select the Committee's members, and weak oversight of the CDF program by both citizens and the national government. This control was diminished in 2013 with the passage of an amended Constituencies Development Fund Act. However, the period we study in this paper was prior to the tightening of these rules.

[3] As indicated in Table 1, an average of 77 percent of the funds allocated were actually spent on projects. Although this implies that some MPs were "leaving money on the table," this figure compares favorably with similar decentralized fiscal transfer schemes in other settings. For example, Chong et al (2015) report that mayors in Mexico spend on average just 56 percent of the Fund for Social Infrastructure transfers they receive from the central government.

The Constituencies Development Fund Act stipulates that CDF funds can be used for any project whose "prospective benefits are available to a widespread cross-section of the inhabitants of a particular area." They therefore tend to be spent on local public goods such as school classroom renovation, bridge repairs, road grading or the construction of local infrastructure such as dispensaries, dip tanks, water projects, public toilets, or police posts.[4] As Table 1 shows, more than half of the projects initiated between 2003 and 2007 were education related, with the next largest categories being water and health. CDF projects are generally completed within one year, although they occasionally stretch over two years or longer.

The CDF program thus presents each MP with a large, annually replenished, exogenously determined amount of money that, subject to minimal restrictions, he may allocate within his constituency with nearly total discretion.[5] This provides the MP with an important tool for alleviating poverty and promoting development. It also provides him with a valuable political resource for rewarding his supporters and/or attracting new backers for his re-election.

Aside from the benefit it provides Kenyan MPs, the CDF program also provides researchers with a nearly ideal opportunity to observe—in 210 separate "laboratories"—whether political actors behave in keeping with our theories of distributive politics. Since, as a local public good, the benefit of a CDF project is inversely correlated with a citizen/voter's distance from it, we can infer which citizens/voters an MP is attempting to aid, reward or attract based on where he chooses to put CDF projects in his constituency.[6] By estimating the association between project placement and levels of political support for the MP in the previous election, we can test whether politicians reward their supporters. And we can do this while controlling for a host of other factors that offer competing explanations for how MPs seek to locate the resources they control.

Three potential objections might be raised about this simple set-up. First, one might wonder whether project allocations can be thought of as independent observations, since placing a project in a particular geographic spot in one period would seem to imply that one cannot place a project in the same spot in the next period. While it is true that the political payoff from placing a second project in an area that has already received one may be less than the payoff from the first project, it is not the case that multiple projects cannot be situated in what, from a geo-coding perspective, might appear to be the same location: one can be a refurbished school classroom; another a clinic; another a

[4] The CDF Act permits a small amount of the funds to be spent on administration and education bursaries, but this is limited to less than a quarter of total expenditure and in practice is much less.

[5] A small caveat regarding the exogeneity of the transfer amount received by each constituency: Although the allocation mechanism would appear to be tamperproof, it does generate incentives for MPs to lobby the Kenya National Bureau of Statistics to revise their reported poverty rates upward, and thus be able to claim larger shares of CDF funding, and Wahome (2008) reports that some MPs have indeed attempted this strategy. We believe that such instances are rare, however.

[6] An alternative (complementary) analysis might weight the projects by the amount of money spent. Here we focus just on the spatial allocation of projects, treating projects of unequal size as providing equivalent utility.

borehole.[7] Furthermore, the need for development projects of the type that CDF underwrites is so great in most of Kenya that receiving one project does not remove the need for others.

Second, one might wonder if it is reasonable to trust the project records we are using in our analyses. If the CDF records indicate that KSh 600,000 was spent on the construction of teachers' quarters and a pit latrine at a primary school, how sure can we be that the money was actually used for this purpose? The honest answer is that we cannot be certain in every case that the CDF records accurately reflect how the money was actually spent. Our analyses focus on the first five years of the CDF program, and financial oversight of CDF funds during this period was weak. Newspaper accounts, investigative reports, and audits by groups like the National Taxpayers Association have identified many examples of ghost projects, unaccounted for money, and general mismanagement of CDF program funds (e.g., Mutoro 2005; Lumwamu and Munene 2006; Awiti 2008; National Anti-Corruption Campaign Steering Committee 2008). However, the sheer numbers of projects we analyze, combined with the fact that reporting inaccuracies are unlikely to be correlated with where projects are located, suggest that while errors in the project data may attenuate any association between project placement and political support they are not likely to introduce significant bias to our estimates. Even if MPs attempted to mask political favoritism by over-reporting projects in opposition areas—perhaps the most likely source of inaccuracy in the figures we use—the resulting bias would lead to an understatement of the relationship between the MP's support and project placement.

A third potential objection is that the intended beneficiaries of the projects we study might not be the communities living adjacent to them but the contractors who were hired to do the work and/or the suppliers who were awarded the tender for providing the materials (who might live in completely different areas). Since our inference about who the MP is targeting depends on our assumption that the beneficiaries of a project are the people living near it, this would seem to be a problem. However, even if an MP's principal objective in using the CDF funds is to channel contracts to his cronies, he still has to determine where to locate the projects, and he would be foolish not to be strategic in making this (perhaps secondary) decision. The only way that pursuing the first strategy (using the project as an opportunity to award a contract) might interrupt the second (using the project as an opportunity to place a valuable public good in the vicinity of a set of voters the MP wants to reward or from whom he wants to win support in the future) is if the type of project that is suitable for contracting leads for some reason to a bias against placing a project in a particular area. But we think this is not particularly likely.

Our Data

MPs are required to submit annual reports to the national CDF Board detailing how they

[7] The ability of projects to be placed seemingly on top of one another is reinforced by the fact that our geo-referencing procedures (described in further detail below) locate projects, at best, at the level of the village.

have used their CDF allocations. These reports, which are posted in pdf form on the CDF Board's web site, provide project names (e.g., Gatina Dispensary; Kipruti Water Project), a description of the activity to be done (e.g., roofing, plastering, painting one classroom; construction of cattle dip; building of foot bridge), details about the sector in which the project is located (e.g., health, education, water), information about the amount of money allocated to the project, and a description of the project's implementation status (e.g., ongoing, complete, not started). Our principal data, a geo-referenced set of 34,139 CDF projects in 204 constituencies for the 2003 to 2007 financial years, are extracted from this source.[8] We downloaded the pdfs—more than 6,000 pages—and scraped the information, either manually or automatically depending on whether the files submitted by the constituency contained embedded text or an image. We then identified individual projects with multiple records (e.g., projects that received allocations across multiple years) and reformatted them to record each project as a single row, with columns for allocations received in each year.

Since project data from the website did not include any explicit geographic coordinates, we geo-referenced the project records by matching project names to facilities for which point or polygon data were available (e.g., schools, market towns, health centers, water/irrigation features).[9] As indicated in Table 2, we were able to match 60 percent of projects to their exact geo-referenced locations. For most of these, we also had information on the administrative location or ward in which the project was located, which further increased our confidence that the project was matched to the correct point in the correct small area. In another 21 percent of cases we were able to match the project to the correct enumeration area (of which there are an average of 32 per constituency). In about 4 percent if cases we were able to match the project to the sub-location (of which there are an average of about 20 per constituency), and in another 6 percent of cases we were able to match the project to its administrative location (of which there are an average of about 12 per constituency). In the remaining 9 percent of cases we were unable to match the project to any point or administrative sub-unit, so we randomly imputed the project to a site somewhere within the constituency. In all cases where projects were not matched to a specific point—that is, for the 40 percent of projects matched to an enumeration area, sub-location, location or to the constituency more broadly—we limited the points to which projects could be randomly assigned to the inhabited areas of those units. Specifically, following Edelsbrunner, Kirkpatrick and Seidel (1983), we calculated the $\alpha-$shape of all infrastructure points and market towns, dilated it, and intersected that polygon with the unit polygon (e.g., the enumeration area, sub-location, etc.) to estimate the inhabited expanse within the unit. In effect, this step prevents us from inadvertently locating projects in national forests, parks, or reserves.

[8] www.cdf.go.ke, accessed May 28, 2012. Information for six constituencies was missing from the data posted on the CDF Board website.

[9] The Kenya Open Data website contains geo-coded CDF data from 2003 to 2010. However, we determined this data to be untrustworthy. First, most of the geo-coordinates that are provided are the centroid of an administrative location (within which the actual project may or may not be situated. The data set contains 40,000 projects but only 8,000 unique geocode points. Second, the list of projects appears to be incomplete, with several constituencies missing, and others containing very few projects. Finally, the provenance of the data is unclear. Hence, we decided to use a different data source compiled from primary documents, and to code the geo-locations of projects ourselves.

Although these geo-referencing procedures are not perfect, they should introduce only noise, not bias, into our estimates. In any event, our results are robust to dropping the 19 percent of projects that are matched with a level of specificity below the enumeration area.[10]

(Table 2 Here)

To estimate the spatial association between project placement and support for the MP in the last election, we combine these CDF data with polling station-level results from Kenya's 2002 parliamentary elections. By linking these results to a geo-referenced polling station dataset, we create rasters identifying the number of votes won by the winning candidate at each point in each constituency. To estimate political support at each raster square, we first generate a Voronoi decomposition of the constituency, which divides the area of the constituency into a number of catchment areas delimiting the areas closest to each polling station. These catchment areas can be thought of as the areas from which each polling station draws its voters.[11] Within each catchment area, we allocate political support to raster squares proportional to the population density within that area. With more than 14,000 polling stations countrywide and an average of roughly 67 polling stations per constituency this procedure generates an extremely fine-grained and accurate constituency-level mapping of the distribution of electoral support for the incumbent MP.

In addition to these two main variables of substantive interest, we include several important control variables in our analyses. We use the high-resolution population density rasters described in Tatem et al. (2007) to control for the distribution of the population across the constituency. Population density may be important since MPs will have strong incentives to put projects where many people live—either because that will allow them to do the most good or because that is where the largest numbers of voters are located.

We also utilize the World Bank/Kenya Ministry of Roads and Public Works dataset described in GoK (2006) to create a raster identifying the distance from each point in each constituency to a paved road. Distance to a paved road may be important insofar as projects located closer to roads may be cheaper to implement. All else equal, we would expect MPs to favor locations that allow them to do the same thing more cheaply.[12]

Although we argued earlier that the need for development projects is great throughout Kenya (and therefore that the allocation of one project to an area would not necessarily reduce the demand for a second), there is nonetheless meaningful variation in need, both across and within constituencies. If CDF funds are allocated with an eye toward alleviating need (either because the MP desires to improve people's wellbeing or because

[10] See Appendix X.

[11] A key assumption of this approach is that individuals register to vote at the polling station closest to their home. Nationally representative surveys of registered voters in 2002 and 2007 by the Institute for Education in Democracy confirm that about 95% of voters register at the polling station closest to their home.

[12] Distance to a paved road is also negatively correlated with population density (r = -.245). Hence we may also expect it to be relevant for the same reasons.

he calculates that the political returns to a project will be greater if people are more desperate), then controlling for the spatial distribution of need will be important. We do this by drawing on data described in Tatem et al (2013) to create constituency-level rasters that depict estimates of the proportion of the population living in poverty in each one square km grid cell.

Given the close association in Kenya between ethnicity and voting behavior (Barkan and Ng'ethe 1998; Gibson and Long 2009), one might think that an MP's ethnic match with voters will already be built into the distribution of his support in the last election. However, it may nonetheless be of interest to test whether projects are disproportionately channeled to the MP's coethnics.[13] To do this, we employ polling station-level estimates of ethnic demographics from Harris (2013) to create a raster identifying the estimated number of the MP's coethnics at each point in the constituency. We also take a second cut at the expectation that MPs will favor their own by creating a raster for each constituency indicating the distance from each point to the MP's home village. Especially in ethnically homogeneous constituencies—quite common in Kenya[14]—the relevant communal distinction may not be between coethnics and non-coethnics (since there are none of the latter) but between members of the MP's family or clan and other members of the broader community. MPs may also seek to target their home village for purely personal, rather than group-regarding, reasons. To the extent that they plan eventually to retire to the village, allocating development projects there will raise their future comfort level.

Figure 1 provides an illustration of what our data look like for an example case: the Rift Valley constituency of Eldoret South. As the Figure makes clear, our estimation of the spatial association between voting patterns and project placement is based on extremely rich data. Although we describe our methodology in more technical detail in the next section, the most intuitive way of thinking about how we generate our estimates is that we superimpose the map of CDF projects on top of each of these other maps (alone or in combination) and calculate the spatial association between project locations and the independent variable(s) in question—and we do this in 204 different constituencies.

(Figure 1 Here)

Modeling the Association Between Project Locations, Votes, and Other Covariates

To estimate the spatial relationship between where CDF projects are located and our explanatory variables of interest, we employ a Poisson point process model (Gatrell 2006; Diggle 2003). The intuition behind how the model works is straightforward. Start with a two-dimensional geographic space. Project onto it an interpolated gradient of some variable of interest (in our case the number of votes won by the MP in the last

[13] For evidence linking ethnicity to patterns of patronage distribution in Kenya, see Burgess et al (2013) and Kramon and Posner (2014).

[14] Forty-three percent of constituencies in Kenya have a level of ethnic diversity less than 0.3 (calculations from data described in Harris (2013)).

election). Then locate the spatial data points of interest (in our case CDF projects) in that space. These data points are then complemented with an arbitrary number of "dummy" points representing locations where projects were not located. In combination, the actual and dummy points are then used to divide the region into random polygons, which contain outcome information about the number of projects in the polygon (a Poisson variable) and covariate information from the rasters themselves. This "pseudolikelihood" approach allows us to estimate the relationship between the variable of interest (votes) and the number of observed points (projects). This can be done either conditional on covariates or without them.[15] Further details are provided in the Appendix.

This approach has two distinct advantages. First, it can account for multiple projects in the neighborhood of a polling station, each at different distances from it and with different interpolated levels of political support at each project point. Second, it allows us to model spatial variation in project placement directly, rather than using some ad hoc aggregation such as administrative borders. The more common practice in studies of distributive politics is to count projects that fall within an administrative area and then look for correlations between the number of projects that are observed and the characteristics of the unit—its population, its dominant ethnic group or partisan faction, the number of votes won by the ruling party, etc. Such an approach forces the researcher to aggregate outcome and explanatory data to that administrative unit, thereby obscuring the more localized relationship between project placement and characteristics that might vary within the unit. While this problem will become less significant the smaller the unit becomes, it will always preferable to avoid it by estimating the association at the level of the project units themselves. In addition, in Kenya, as in many other places, administrative boundaries are products of political processes (Oucho 2002; Kasara 2006). Using such boundaries to delineate units of analysis could therefore introduce endogeneity, as the boundaries themselves may be products of the very political processes being studied.

What Factors Affect Where CDF Projects are Placed?

The fact that our data and methodology put us in a position to estimate the association between project placement and a set of relevant independent variables in 204 separate constituencies presents challenges for how to present our findings. One possible approach would be to estimate a pooled model across all 204 constituencies and report the results of that single analysis. Unfortunately, the pseudo-maximum likelihood (PML) approach we use is computationally intensive and infeasible with the number of projects we analyze. Also, the way the PML model weights observations and discretizes continuous two-dimensional space depends considerably on the sizes of the units, and the marked variation in Kenya in the sizes of constituencies generates complications. Large constituencies would contribute large amounts of information to the likelihood, while smaller constituencies less.

[15] We estimate all models using the spatstat package in R (Baddeley and Turner 2005).

There are also theoretical considerations. A pooled regression assumes that the process under investigation (project placement) happens the same way in all units, and can thus be estimated appropriately in a single regression equation. This is at odds with both the evidence we present below and our priors, which are that MPs face different electoral contexts and local conditions and will thus pursue different strategies of project allocation. We therefore eschew a pooled analysis and instead present the distribution of constituency-level estimates for each variable of interest in a series of boxplots (see Figure 2).[16] A drawback of this choice is that the boxplots do not provide information about the statistical significance of each constituency-level estimate. But they do provide a clear way of depicting the general pattern of relationships across the whole of the country. The first six columns of Figure 2 show bivariate relationships between project placement and the variable in question; the final column shows the relationship between votes and project placement conditional on all of the other independent variables.[17]

(Figure 2 Here)

As expected, there is a positive association in nearly all constituencies between population density and project placement. Either because they want to benefit the most people or because they want to win the most votes, most MPs do seem to allocate CDF projects to areas where the largest numbers of people are located. Distance from paved roads also has the expected negative sign in most constituencies, suggesting that MPs tend to put projects closer to (that is, at a smaller distance from) roads.

Poverty is generally not positively associated with project placement. Indeed, in most constituencies projects tend to be placed in areas that are less poor. One implication of this finding is that, contrary to the CDF program's stated objectives, CDF funds are not being used as a tool for poverty alleviation—or at any rate are not being channeled to the areas with the largest proportion of poor people.[18] Another interpretation is that poverty cuts two ways. On the one hand, it makes people more needy (and, from a social planner's perspective, more deserving of development resources). But on the other hand, it makes people less able to mobilize to demand that projects be located in their areas. Recall that while CDF allocation decisions are made by the MP, community members may also (and frequently do) apply for projects. The negative association between poverty and project placement in many constituencies may simply reflect the weakness of poor people in making such demands.[19]

[16] In all estimates, we standardize the measure for cross-constituency comparability by subtracting the mean and dividing by the standard deviation.

[17] We drop population density because it is highly correlated with the number of votes won by the MP (mean correlation across the 204 constituencies of .633).

[18] Note, however, that our measure of poverty captures only the rate of poverty in a location, not its depth. So it is still possible that MPs are channeling projects to the very poor—providing that they live in isolation from larger concentrations of less poor people.

[19] If poverty rates are higher in rural areas, then the result on our poverty measure is consonant with the population density findings. Poverty may also be negatively associated with turnout, which means that, all else equal, a vote-maximizing MP would be less inclined to locate projects in the poorest areas.

The vast majority of MPs appear not to locate projects in areas dominated by their coethnics. Although perhaps surprising given the centrality of ethnicity in Kenyan politics, this is likely because most Kenyan electoral constituencies are ethnically homogeneous, so targeting projects along ethnic lines is not possible. We do find evidence, however that MPs are more likely to locate CDF projects in proximity to their home village. As noted earlier this could be because the MP's home village is where his family and/or clan is located, and family/clan loyalties are salient in a context where everyone in the constituency belongs to the same broader ethnic category.

We turn now to our main relationship of interest: the association between project placement and the number of votes won by the MP in the last election. In both the bivariate relationship (column 6) and when we also control for distance to roads, poverty, coethnicity with the MP, and distance from the MP's village (column 7), we find that in the vast majority of constituencies CDF projects are indeed more likely to be placed in areas where the MP won a large number of votes.[20] These findings suggest that Kenyan MPs do in fact favor their supporters in the allocation of CDF projects.

These general patterns are not, however, universal. As the boxplots make clear, there is significant cross-constituency variation in the association between project placement and each explanatory variable. In the case of every variable we study, constituencies in which the association with project placement is positive are counterbalanced by constituencies in which the association is negative—and, of course there are many in which we would conclude there is no association once we accounted for statistical significance. So the key question for each variable may not be "how does this factor affect the way MPs allocate CDF projects?" but "why does this factor affect the way MPs allocate CDF projects one way in one constituency and different way in another constituency (and not at all in yet another constituency)?"

We further explore this heterogeneity for our main quantity of interest in Figure 3. The Figure plots the coefficient estimate and 95 percent confidence interval for the conditional association between project placement and votes won by incumbent in the previous election in each of the 204 constituencies. While the Figure confirms the general positive relationship between these variables (the error bars are, in the vast majority of cases, to the right of the zero effect line), it also demonstrates clearly that the relationship varies considerably across constituencies. In places like Belgut in Kericho district, Changamwe in Mombasa district, Kasarani and Langata in Nairobi district, Manyatta in Embu district, and Mathira in Nyeri district (to select but a few of the constituencies with large positive coefficient estimates) our results suggest that MPs strongly steered projects to their supporters. But in places like Kasipul-Kabondo in Rachuonyo district, Kiharu in Muranga district, Makadara and Starehe in Nairobi district, and Webuye in Bungoma district (to select but a few of the constituencies with well-estimated zero coefficients) we find no evidence that MPs were any more likely to channel CDF projects to the people who voted for them in the last election than to the people who supported other candidates. The Figure suggests that the really interesting

[20] In generating these estimates, we take the log of the number of votes won by the MP prior to standardizing the measure.

question may not be whether politicians reward their supporters but *under what conditions*.

(Figure 3 Here)

Under What Conditions do MPs Favor Their Supporters?

Answering this question requires a more nuanced understanding than we have provided thus far of the incentives and constraints that MPs face in the context we study. The most relevant fact facing MPs in Kenya is that re-election rates are not high: among candidates who won in 2002 and ran again in 2007, just 40 percent were re-elected.[21] This implies that MPs should be strongly incentivized to use whatever resources they have at their disposal to increase their chances of securing another term.[22] CDF projects are tremendously useful for this purpose, since the principal output that Kenyan voters look to their MPs to provide is constituency service and local public goods (Barkan et al 2010).

The optimal strategy for an MP is to locate projects in the vicinity of large numbers of voters, and particularly those whose voting decisions are likely to be swayed by the provision of the project. This implies that MPs should prioritize areas containing large numbers of people who voted for them in the last election, since these are the people most likely to support them again conditional on receiving a reward for having voted for them in the past. If the MP was elected by a very large margin, and if past supporters can in fact be convinced to support the MP again if they are given CDF projects, then following this strategy may be enough to win again. In such a situation we would expect MPs to concentrate their CDF spending on areas where they received high levels of support in the last election and to ignore areas with few supporters.

But if the MP's margin of victory in the last election was low, he may not be able to feel confident about his re-election if he focuses only on his past supporters. In such a situation, he will have incentives to expand the distribution of CDF projects to areas

[21] Of the 209 constituencies for which parliamentary election results are available for both years, there are 175 cases in which the candidate who won in 2002 ran again in 2007 (whether incumbents who did not run for re-election died in office or simply chose not to run again is not clear from these data). Of these, just 70 retained their seats. One reason for this low re-election rate is because of the significant realignment in the party system that took place in Kenya between 2002 and 2007. Voters in Kenya put more weight on party tickets than on the candidates themselves in multiparty parliamentary elections (Posner 2007), so many of the incumbents who lost may simply have bet wrong on the party with which to affiliate themselves in 2007 (or they got lucky in 2002 by being affiliated with what, at that time and for the part of the country in which they were running, was the right party, and then revealed themselves not to be very good MPs). Even so, the low re-election rate underscores the extent to which most MPs cannot take re-election for granted.

[22] Of course it is possible that, understanding how difficult the task of getting re-elected may be, MPs simply accept that they are not likely to win again and thus ignore strategic considerations altogether. But this is belied by the fact that 84 percent of the candidates who won in 2002 ran for re-election in 2007—no doubt a lower bound since this includes (an undetermined number of) those who had died in office and therefore could not run.

containing people who did not vote for him in the past but who might be convinced to do so in the future given the proper inducements.[23] The question then becomes: which voters might be converted? In the Kenyan context, a plausible answer is that non-coethnics are not likely to be convertible but coethnics who voted for a different candidate in the last election might be. Hence, if the MP is in a constituency where nearly all non-supporters are non-coethnics, he will have little incentive to expand his distribution of CDF projects to areas that do not contain large numbers of voters who supported him in the last election. But if he is in a constituency where non-supporters are coethnics, it may be possible to win them over by locating projects in their areas. We would therefore expect to see a mixed strategy, with some projects being directed toward past supporters and some projects being directed toward past opponents.

The spatial implications of such a strategy, however, depend on whether or not supporters and non-supporters are geographically segregated from one another. If they are segregated, then allocating CDF projects to both supporters and past opponents will generate a weaker association between votes won in the past election and project placement. But if supporters and non-supporters are intermixed with one another (i.e., not segregated), then the benefits of projects targeted at supporters will spill over to non-supporters. This means that the MP can accomplish the goal of reaching across to non-supporters by continuing to channel projects to the people who voted for him in the last election.

This logic, as well as its implications for our estimates of the association between MP votes and project placement, is summarized in Figure 4. In constituencies where the MP's margin of victory in 2002 was high, he can rely on his own supporters to win re-election, so we would expect a strong association between votes won and project placement. When the margin of victory was low, the expected outcome depends on whether the MP faces non-supporters who are members of his own ethnic group. Although it is difficult to specify ex ante when this will be the case, the ethnic heterogeneity of the constituency serves as a reasonable proxy. In effect, this measure sorts constituencies into those where nearly everyone is a member of the same group (and the MP's opponents will be co-ethnics by definition) and those that are ethnically mixed, where patterns of electoral support tend to follow ethnic lines and the MP's opponents are likely to be non-coethnics. In the latter case, we would expect to see a strong association between project placement and past electoral support, since the conversion of non-supporters to supporters will be unlikely.

(Figure 4 Here)

In an ethnically homogeneous setting, however, the expected association will depend on the segregation of past supporters and non-supporters. Where they are intermixed, spillover will be high and MPs will be able to simultaneously court past supporters and

[23] The MP might also pursue a "turnout buying" strategy by channeling projects to areas containing large numbers of people who did not vote in the last election but who might be convinced to vote next time (and to support him) if they are given projects. We leave this possibility aside for the moment, but will address it in future iterations of the paper.

non-supporters by locating projects near past supporters. This should generate a strong association between votes and project placement. Where past supporters and non-supporters live apart from one another, spillover will be low and the attempt to reach out to non-supporters will weaken the association between past votes and projects.

In explaining the variation we find in how much MPs reward their supporters with CDF projects, the MP's vote margin in the last election, the constituency's ethnic heterogeneity, and the extent to which the MP's supporters and opponents are physically segregated are likely to be important factors. But these are not the only factors that may matter. For example, there is evidence that patterns of public goods provision may vary with the politician's gender (Chattopadhyay and Duflo 2004). If male and female MPs make different calculations in weighing the tradeoff between channeling CDF projects to those with the greatest need and those with the greatest political payoff, we might expect to find different associations between votes and project placement across MPs of different genders.

Whether or not an MP is a member of the ruling coalition may also matter. While CDF funding represents a considerable source of capital for local public goods, it is not the only source. Central government ministries in Kenya also spend millions of dollars a year on roads, schools, health facilities, and other local infrastructure. To the extent that MPs with ties to the ruling coalition have greater access to (or greater control over the targeting of) such resources, they may be less dependent on CDF funds and more free to locate CDF projects in less politically strategic places. Hence we might expect to find a weaker association between votes won and project placement in constituencies where the MP is a member of the ruling coalition.

Insofar as knowing how to deploy resources to maximize one's re-election prospects requires skill and political experience, an MP's status as a neophyte versus a long-time incumbent is likely to be relevant. We might therefore expect to find the association between votes and projects to be increasing in the number of elections that the MP had won prior to 2002.

Another possibly relevant factor—this time a characteristic of the constituency rather than the MP—is the extent to which the population in the constituency is mobilized politically. When they are (as indicated by high election turnout), MPs may need to be especially attentive to their constituents' needs and to behave more strictly in keeping with their optimal political strategy. We might therefore expect to see turnout rates in 2002 be associated with a stronger association between votes and project placements.

Finally, it is plausible that we may find different patterns in urban and rural constituencies—a distinction we can proxy for through (the log of) population density.

We test these expectations in Table 3.[24] We first focus on the three main variables of interest as outlined in Figure 4: margin of victory, ethnic heterogeneity, and segregation

[24] Note that the analysis and discussion here is very preliminary, so comments are most welcome!

of MP's support.[25] When entered into a regression equation singly, as in columns 1-3, all three variables have positive coefficients, though only ethnic heterogeneity is statistically significant. Column 4 includes all three variables together, and we see both the magnitude and statistical significance strengthen, suggesting that all three factors relate positively to strengthened targeting. The positive coefficient on victory margin is consistent with the first node of Figure 4, as higher victory margins correspond to more single-minded targeting of supporters as the incumbent MP's electoral cushion allows him to ignore constituents who did not support him in the last election.

Column 5 includes the other variables discussed above: whether the MP is female, whether the MP is a member of the ruling coalition, prior political experience, turnout in the past election, and population density.[26] None can be differentiated from zero, and all point estimates are close to zero. This comparison suggests that our general ideas about how the spatial structure of the electorate (segregation/integration), the safety of the seat (vote margin), and the pliability of political preferences (ethnic heterogeneity) have more explanatory power than characteristics of the MP, turnout, or population density.

To better capture the logic of what happens lower down in Figure 4 (where the expected signs are conditional on the signs of the variables higher up in the logic tree), Columns 6 to 8 examine interaction effects on the three main terms of interest. Column 6 shows a positive coefficient on the interaction between ethnic heterogeneity and victory margin, suggesting that each variable's relationship with targeting strength is dependent on the value of the other. Holding ethnic heterogeneity at its mean implies that a one unit change in victory margin corresponds to a 0.26 unit strengthening in targeting. Similarly, at the average victory margin, a one unit increase in ethnic heterogeneity corresponds to a 0.11 unit strengthening in targeting. Substantively, the interaction of heterogeneity and victory margin amplifies the propensity of an MP to target his supporters: when both are high, we see a stronger association between votes in the last period and project placement. This is consistent with the logic outlined in Figure 4.

Column 7 examines the interaction between victory margin and segregation, and finds a positive but insignificant interaction. Column 8 shows the interaction of segregation and ethnic heterogeneity. The interpretation of the negative coefficient is that when larger numbers of non-supporters live amongst supporters (i.e., when segregation is low), the relationship between ethnic heterogeneity and targeting becomes weaker. Conversely, when supporters and non-supporters live in distinct areas (i.e., segregation is high), the relationship between ethnic heterogeneity and targeting becomes stronger. Put another way, the more segregated supporters are from non-supporters, targeting of supporters strengthens as a function of ethnic heterogeneity. Similarly, in more homogeneous

[25] Following Reardon and O'Sullivan (2004), we define segregation in terms of the spatial exposure of group n in the local environments of each member of group m.

$$ {}_m\tilde{P}^*_n = \int_{q \in R} \frac{\tau_{qm}}{T_m} \tilde{\pi}_{qn} dq $$

[26] Prior political experience is proxied as the number of elections won by the MP between 1992 and 2002. Since there were just two elections during that period, it therefore takes values of 0, 1, and 2.

constituencies, MPs are more apt to reward supporters even when spatially integrated with non-supporters. This is consonant with the middle part of Figure 4. In ethnically heterogeneous constituencies, non-coethnic voters cannot be swayed, and thus MPs target (likely coethnic) supporters.

These results, though still tentative at this stage, are broadly consistent with the theory and underscore the important (but, in the distributive politics literature, usually overlooked) importance of spatial features in determining the strength of partisan targeting.

Conclusion

TO BE WRITTEN...

References

Awiti, Victoria Phildah. 2008. "An Assessment of the Use and Management of Development Funds: The Case of Constituencies Development Fund in Kenya." Masters Thesis.

Baddeley, Adrian and Rolf Turner. 2000. "Practical Maximum Pseudolikelihood for Spatial Point Patterns." *Australian and New Zealand Journal of Statistics* 42(3):283-322.

Baddeley, Adrian and Rolf Turner. 2005. "Spatstat: An R Package for Analyzing Spatial Point Patterns." *Journal of Statistical Software* 12(6):1-42.

Barkan, Joel and Michael Chege. 1989. "Decentralizing the State: District Focus and the Politics of Reallocation in Kenya." *Journal of Modern African Studies* 27, 3: 431-453.

Barkan, Joel and Njuguna Ng'ethe. 1998. "Kenya Tries Again." *Journal of Democracy* 9 (April):32-48.

Barkan, Joel, Robert Mattes, Shaheen Mozaffar, and Kimberly Smiddy. 2010. "The African Legislatures Project: First Findings." *Centre for Social Science Research Working Paper No. 277*.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padro i Miquel. 2013. "The Value of Democracy: Evidence form Road Building in Kenya." Unpublished ms.

Chattopadhyay, Raghabendra and Esther Duflo. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India." *Econometrica* 77(5): 1409-1443.

Chong, Alberto, Anna De La O, Dean Karlan and Leonard Wantchekon. 2015. "Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice and Party Identification." *Journal of Politics* 77(1): xx-xx.

Cox, Gary. 2006. INSERT SOURCE

Cox, Gary and Matthew McCubbins. 1986. "Electoral Politics as a Redistributive Game." *Journal of Politics* 48(2): 370-389.

Diggle, Peter J. 2003. *Statistical Analysis of Spatial Point Patterns*. London: Arnold.

Dixit, Avinash and John Londregan. 1996. "The Determinants of Success of Special

Interests in Redistributive Politics." *Journal of Politics* 58(4): 1132-1155.

Edelsbrunner, H., D.G. Kirkpatrick and R. Seidel. 1983. "On the shape of a set of points in the plane." *IEEE Transactions on Information Theory* 29(4):551-559.

Franck, Raphael, and Ilia Rainer. 2012. "Does the Leader's Ethnicity Matter? Ethnic Favoritism, Education and Health in Sub-Saharan Africa." *American Political Science Review* 106(2): 294-325.

Gibson, Clark C. and James D. Long. 2009. "The Presidential and Parliamentary Elections in Kenya, December 2007." *Electoral Studies* 28(3): 492-517.

GoK. 2003. "The Constituency Development Fund Act, 2003." Nairobi: Government Printer.

GoK. 2006. "Classified Roads and Market Town Datasets." Technical Report. Nairobi: Ministry of Roads and World Bank.

Harris, J. Andrew. 2013. "A Method for Extracting Information About Ethnicity from Proper Names." Unpublished paper.

Institute of Economic Affairs. 2006. "Kenya's Verdict: A Citizens Report Card on the Constituencies Development Fund (CDF). IEA Research Paper Series, No. 7.

Jablonski, Ryan. 2014. "How Aid Targets Votes: The Impact of Electoral Incentives on Foreign Aid Distribution." *World Politics* 66(2): 293-330.

Kasara, Kimuli. 2006. "Ethnic Beachheads and Vote Buying: Explaining the Creation of New Administrative Districts in Kenya, 1963-2001." Unpublished paper.

Kramon, Eric and Daniel N. Posner. 2014. "Ethnic Favoritism in Primary Education in Kenya." Unpublished paper.

Lumwamu, Kennedy and George Munene. 2006. "3 Charged with Theft of CDF Cash." *Daily Nation*, 5 December.

Morjaria, Ameet. 2011. "Electoral Politics and Deforestation." Unpublished ms.

Mutoro, Stephen. 2005. "CDF Has Become a Milch Cow for MPs." *Daily Nation*, 30 March.

National Anti-Corruption Campaign Steering Committee. 2008. "The Constituency Development Fund: An Examination of Legal, Structural, Management and Corruption Issues in Kenya."

Oucho, John. 2002. Undercurrents of Ethnic Conflict in Kenya. Leiden: Brill.

Posner, Daniel N. 2007. “Regime Change and Ethnic Cleavages in Africa.” *Comparative Political Studies* 40, 11 (November): 1302-1327.

Reardon, Sean F. and David O’Sullivan. 2004. “Measures of Spatial Segregation.” *Sociological Methodology* 34(1): 121-162.

Tatem, Andrew J., Abdisalan M. Noor, Craig von Hagen, Antonio Di Gregorio and Simon I. Hay. 2007. “High Resolution Population Maps for Low Income Nations: Combining Land Cover and Census in East Africa.” *PLoS ONE* 2(12):e1298.

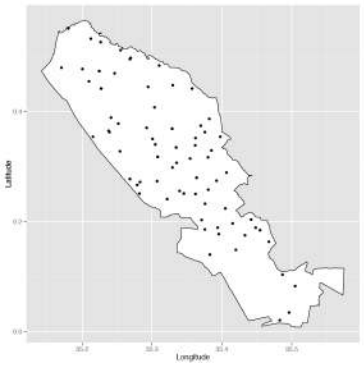
Tatem, A.J., P.W. Gething, S. Bhatt, D. Weiss and C. Pezzulo. 2013. “Pilot High Resolution Poverty Maps.” University of Southampton/Oxford.

Wahome, Mwaniki. 2008. “Study Finds Fault in Allocation Criteria for CDF.” *Daily Nation*, 11 October.

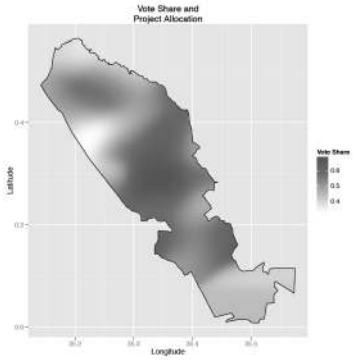
Warton, David and Leah Shepherd. 2010. “Poisson Point Process Models Solve the “Pseudo-Absence problem” for Presence-Only Data in Ecology.” *Annals of Applied Statistics* 4(3):1382-1402.

Wrong, Michela. 2009. *It’s Our Turn to Eat. The Story of a Kenyan Whistle-Blower.* New York: Harper Collins.

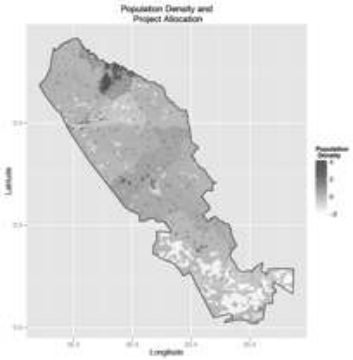
Figure 1: An Illustration of our Data in a Sample Constituency



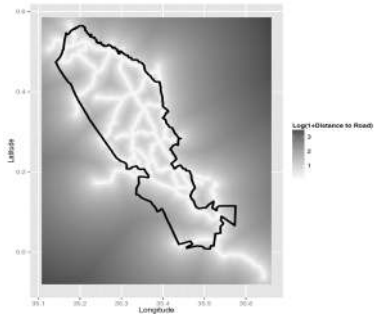
Projects



Votes won by MP in 2002

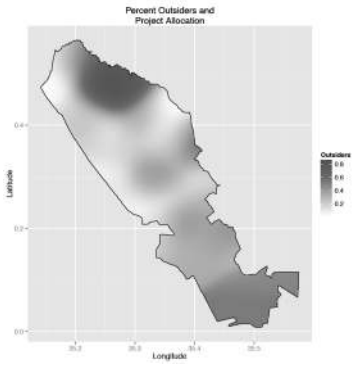


Population Density



Distance to Roads

% Living in Poverty



MP's Ethnic Group

Distance to MP's Village

Figure 2: What Factors Affect Where CDF Projects Are Placed?

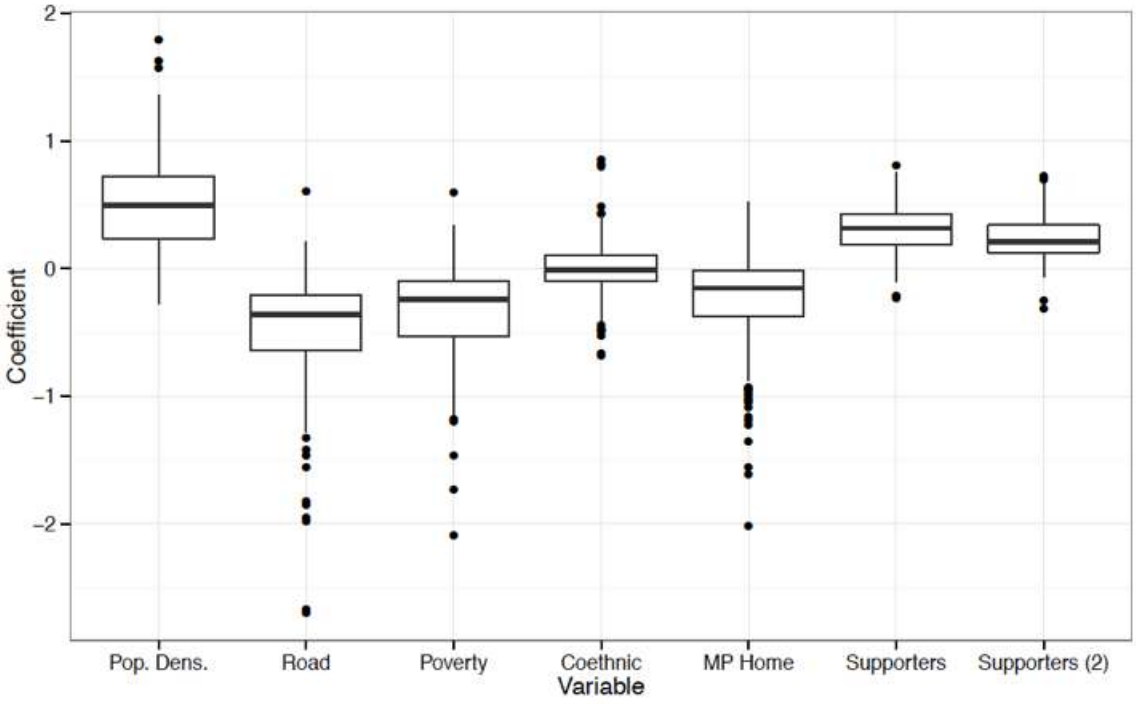


Figure 3: Association Between Votes Won by the MP and Project Placement, by Constituency

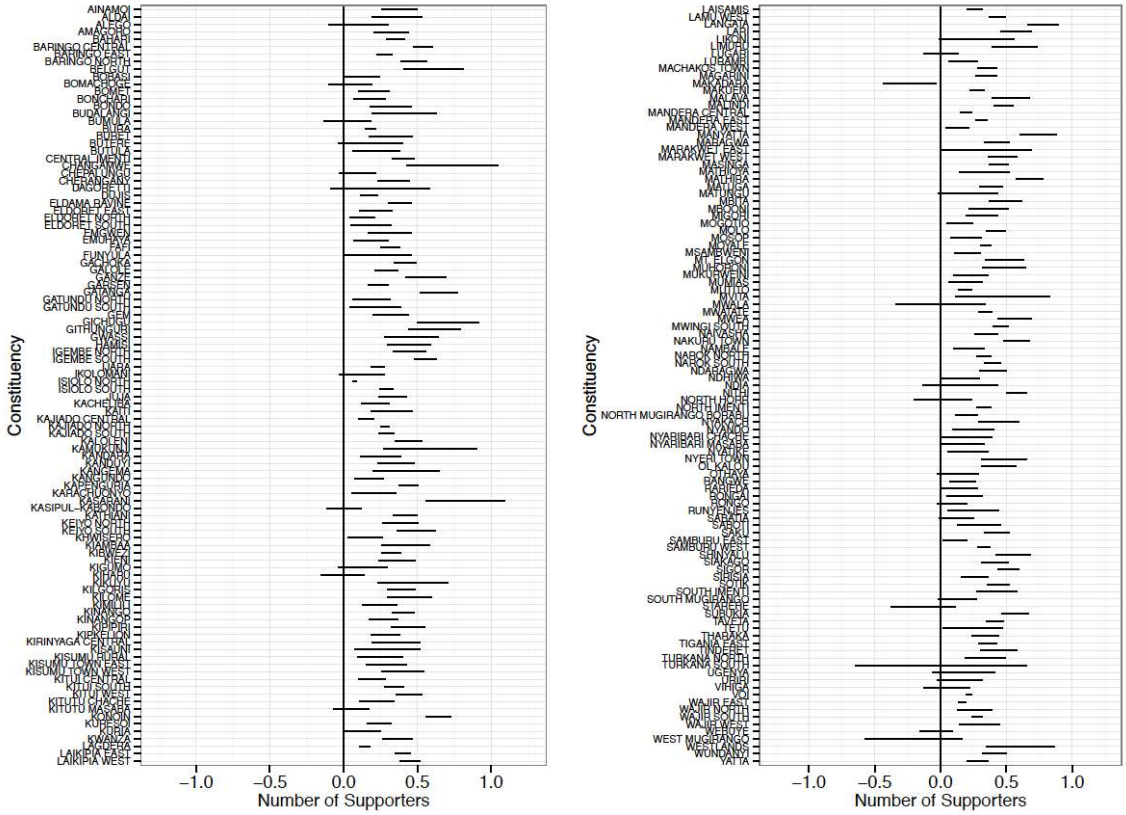


Figure 4: Expected Strength of Association Between MP Votes and Project Placement

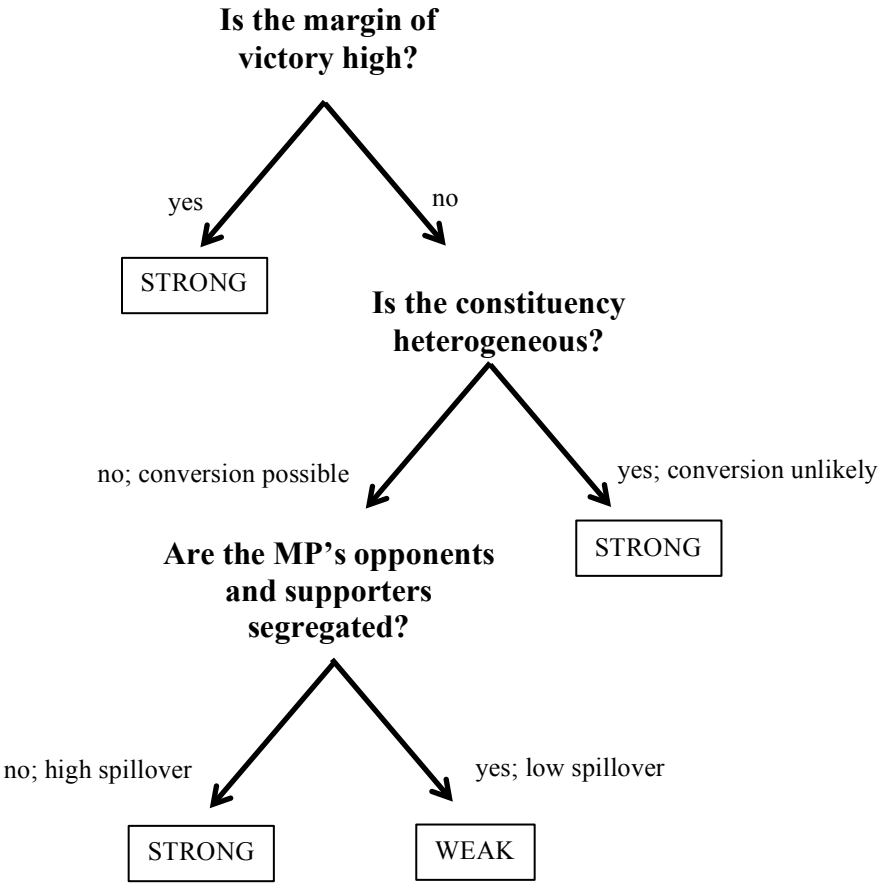


Table 1: Basic Statistics on CDF Projects in Kenya, 2003-2007

CDF program funding	
Total allocations[1]	US$ 369,494,222
Avg allocation per constituency	US$ 351,900 (min=$66,667; max=$675,401)
Avg % of allocation spent	77
Projects	
Total number of projects	34,139
Avg per constituency	163 (min=14; max=589)
Avg % allocated to 3 largest projects	11
Avg Herfindahl index of project spending	0.08
Project types (%)	
Health	9
School	54
Water	13
Other	15
Project duration (%)	
1 year	57
2 years	25
3 years	12
4 years	5
5 years	1

[1] 1 US$ = 90 KSh. Less than one percent of projects were labeled as bursary or administrative projects and are thus excluded from this table. Turkana South constituency reported the minimum number of projects (14); Eldoret North reported the maximum (589). Funyula constituency reported only 8 projects, as recorded in our database, but this is due to an inability to disambiguate individual projects implemented within the 8 geographic subunits of that constituency. Due to errors in the source data, some constituencies showed allocations to projects totaling over 100% of their constituency-level allocations for a given year. These are assumed to be equal to 100%.

Table 2: Matching of CDF Projects

Projects matched to...	number	%
Exact location	20,521	60.1
EA	7,030	20.6
Sub-location	1,311	3.8
Location	2,146	6.3
Constituency	3,131	9.2
Total	34,139	100

Table 3: Under What Conditions do MPs Favor their Supporters?

	Dependent variable:							
	Association between # votes won by MP and Project Placement							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
margin of victory in 2002	0.04			0.27***	0.27**	0.13	0.18	0.26***
	(0.05)			(0.09)	(0.11)	(0.12)	(0.12)	(0.09)
ethnic heterogeneity		0.08*		0.10**	0.10**	-0.02	0.10**	0.30***
		(0.04)		(0.04)	(0.05)	(0.07)	(0.04)	(0.11)
segregation of MP's support			0.07	0.80***	0.87**	0.76**	0.61*	1.11***
			(0.15)	(0.30)	(0.35)	(0.30)	(0.34)	(0.34)
female					0.02			
					(0.07)			
member of ruling coalition					0.03			
					(0.03)			
# of elections won 1992-2002					-0.005			
					(0.03)			
% turnout in 2002					-0.05			
					(0.17)			
log population density					-0.002			
					(0.01)			
margin of victory in 2002 * ethnic heterogeneity						0.36*		
						(0.18)		
margin of victory in 2002 * segregation of MP's support							0.79	
							(0.63)	
segregation of MP's support * ethnic heterogeneity								-1.05*
								(0.55)
constant	0.30***	0.28***	0.30***	0.03	0.05	0.09	0.06	-0.02
	(0.02)	(0.02)	(0.03)	(0.09)	(0.12)	(0.09)	(0.09)	(0.09)
observations	204	204	204	204	204	204	204	204
R^2	0.003	0.02	0.001	0.06	0.06	0.07	0.06	0.07
adjusted R^2	-0.001	0.01	-0.004	0.04	0.03	0.05	0.04	0.05

Note: *p<0.1; **p<0.05; ***p<0.01

Appendix

Following Baddeley (2010), the observed data are the locations of *n* development projects, $\mathrm{x} = \{x_1, \ldots, x_n\}$. This point pattern is a realization of the point process X in a given constituency W; $\mathrm{x} \in W$. The Poisson process model described here focuses on estimating parameters of the *intensity function* for all locations $u \in W$. Let $E[N(\mathrm{X} \cap B)]$ be the expected number of points in B, a region within W. The intensity function is:

$$E[N(\mathrm{X} \cap B)] = \int_B \lambda(u)du$$

For W, we can estimate the intensity as the count of points in x divided by the area of W. This is the intensity in the entire constituency. Point patterns may not occur with uniform intensity, since some areas of a constituency likely receive more projects than others.

$\lambda(u)$ is the intensity of a local Poisson process at location u. Note that covariates Z are measured at every point in the window W. The stochastic component of the model is defined as:

$$\mathrm{X} \sim \mathrm{Poisson}\big(\lambda(u)\big)$$

The systematic component of the model is defined as:

$$\lambda(u) = e^{Z(u)\beta}$$

$Z(u)$ are the values of spatial covariates at location u.

Baddeley and Turner (2000) gives an accessible introduction to estimation. This approach partitions the continuous spatial domain of interest (e.g., the constituency) by adding arbitrary "dummy" points to the realized points where projects are placed. Results presented in the paper are generally robust to changes in the number and location of dummy points used during estimation. Intuitively, this approach can also be thought of as adding "pseudoabsences" (e.g., locations of places where projects were not placed) to estimate parameters from the log-likelihood (Warton and Shepherd 2010).