

**(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

16 March 2015

*Work in Progress:*

*Please Contact Authors for Up-to-date Draft Before Quoting or Citing*

\*Prepared for the UC Berkeley Comparative Politics Colloquium, 19 March 2015.

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 provide evidence for this pattern (e.g., Barkan and Chege 1989; Franck and Rainer 2012; Morjaria 2013; Burgess et al 2014; Kramon and Posner 2014; Jablonski 2014). However, a weakness of this literature is that it focuses exclusively on the allocation decisions of national-level political actors, ignoring the strategies pursued by elected representatives seeking to secure their political futures within their own electoral constituencies.<sup>1</sup> In addition, most of these studies employ data on allocations that are aggregated to the district (or, occasionally, the constituency) level.<sup>2</sup> This paper offers an advance over prior work 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 project level. From a theoretical standpoint 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. From an empirical standpoint it represents an improvement over earlier work by permitting both more precise inferences about who benefits from the allocation of resources and an analysis of how the spatial distribution of political support affects how resources are targeted.

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 on the precise geo-locations 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 more than 200 constituencies. Taken together, these data and methods put us in a unique position to learn whether politicians use the resources at their disposal to reward their supporters, and how the extent to which they do this varies 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

---

<sup>1</sup> Hoffmann et al (2014), which focuses on the allocation decisions of elected county councilors, is an exception. However, that paper deals with the allocation of an artificial (and modest) program benefit

<sup>2</sup> Studies that employ DHS data, such as Franck and Rainer (2012) or Kramon and Posner (2014), can identify beneficiaries at the household level. But because DHS data measures outcomes rather than flows of resources, such studies are only able to measure distributive targeting indirectly.

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 and the level of turnout 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.

This paper makes three distinct contributions. First, it is the first attempt of which we are aware to study the allocation decisions of African politicians at the sub-constituency level using project-level data. As noted, existing analyses of distributive politics in Africa focus on district- (or, occasionally, constituency-) level allocations. In doing so, they implicitly assume that all the constituents in a district (constituency) have equal access to the “pork” that is funneled to the unit. 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 or a road accrues primarily to the people living near such infrastructure. Simply counting the number of schools or boreholes in a district may not tell us very much about “who gets what.” For instance Burgess et al (2013) examine the building of paved roads across 41 Kenyan districts between 1963 and 2011. The assumption embedded in the analysis is that roads benefit all of the individuals in a given district equally, rather than those proximate to the road or the communities that the road itself links. Similarly, Morjaria (2013) explores whether districts with political connections to the ruling party suffer higher rates of deforestation, which he interprets as a measure of the government's acquiescence in illegal land clearing by local residents. The assumption, again, is that the individuals who are benefiting from access to the government forestlands are supporters of the ruling party. Unfortunately, without finer grained data, we cannot be sure in either study whether the government's supporters are, in fact, benefiting disproportionately. By geo-referencing every CDF project, our approach allows us to directly examine the spatial allocation of development resources and thus circumvent this ecological inference problem.<sup>3</sup>

Second, by studying the allocation decisions of constituency- rather than national-level actors, we can exploit cross-unit variation in the association we find between project placement and political support to identify the conditions that make it more or less likely that politicians reward their supporters. This represents a major advance over prior work in which a single actor (the president, “the government,” “the ruling party”), subject to a fixed set of conditions, makes the allocation decision. Our analysis, by contrast, explores the allocation decisions of more than 200 separate actors, each operating in a different social, demographic, and political environment.<sup>4</sup>

---

<sup>3</sup> As a secondary benefit, 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.

<sup>4</sup> Studies that exploit changes in regime type or political leaders—for example, Franck and Rainer (2012); Burgess et al (2013); Morjaria (2013); Kramon and Posner (2014); and Jablonski (2014)—effectively study either multiple actors or the same actors in multiple situations. But the number of actors and settings in which patterns of allocation are being compared is tiny compared to the present study.

Finally, our geo-coded project data and high resolution data on voting patterns and demographics make it possible to study how the spatial concentration of a politician's electoral support, as well as the extent to which his supporters and opponents are segregated from one another, shape the politician's allocation strategies. This adds a novel—and, as we demonstrate, highly significant—pair of explanatory factors to the set of usually studied variables, and thus deepens our understanding of politicians' strategic behavior.

### **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 distributed among the country's 210 electoral constituencies according to a formula by which 75 percent of the monies 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. 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. The MP's ability to control the CDF Committee stems both from his statutory role as its chairman and from his ability to select the Committee's members, as well as from weak oversight of the CDF program by both citizens and the national government.<sup>5</sup> As a consequence, 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).

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. Countrywide, these allocations funded a total of 34,139 projects over this five-year period, or an average of about 163 projects per constituency.<sup>6</sup>

(Table 1 Here)

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.” This requirement means that they are spent on local public goods: the rehabilitation of school classrooms, bridge repairs, road grading, water

---

<sup>5</sup> 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.

<sup>6</sup> 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” (or, in Keefer and Khemani's (2009) phrase, “passing on pork”), 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.

projects, or the construction of dispensaries, dip tanks, public toilets, or police posts. 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. 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.<sup>7</sup> This provides the MP with an important tool for alleviating poverty and promoting development. It also provides the MP 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 provides researchers with a nearly ideal opportunity to observe—in more than 200 separate “laboratories”—whether political actors in Africa behave in keeping the conventional wisdom, and with our theoretical priors. Since, as a local public good, the benefit of a CDF project is inversely correlated with a voter’s distance from it, we can infer which voters an MP is attempting to aid, reward or attract based on where in his constituency he chooses to put each CDF project.<sup>8</sup> By estimating the spatial association between project placement and levels of political support for the MP in the previous election, we can test whether politicians reward their supporters.<sup>9</sup> 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.

Before progressing, let’s take a moment to deal with three potential objections to this simple set-up. First, 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 a period when the CDF program was brand new and financial oversight of CDF funds 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

---

<sup>7</sup> A small caveat regarding the exogeneity of the transfer amount received by each constituency: although the formulaic allocation mechanism would appear to be tamperproof, it does generate incentives for MPs to lobby the Kenya National Bureau of Statistics to revise their constituency’s poverty rate upward, and thus be able to claim a larger share of CDF funding. Wahome (2008) reports that some MPs have indeed attempted this strategy. We believe that such instances are rare, however.

<sup>8</sup> The geographic scope of the benefit from a given project may vary with its type—for example, a bridge may benefit people from a wider radius than a borehole. However, all projects should provide utility to people in inverse proportion to their distance from it.

<sup>9</sup> A 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

of CDF program funds during this period (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 the association between project placement and political support they are not likely to introduce significant bias to our estimates.

Second, 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 borehole. 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. 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 (and demand) for others.

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 or the suppliers who were awarded the tender for providing the materials, who might live in completely different parts of the constituency, or even outside the constituency altogether. 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 is to use her CDF funds to channel contracts to her cronies, she still has to determine where to locate the projects, and she 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 affect 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 she 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. We think this is not particularly likely.

## **Our Data**

MPs are required to submit annual reports to the national CDF Board detailing how they 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.<sup>10</sup> 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 MP contained embedded text or an image. We then identified individual projects with multiple records (that is, projects that received allocations across multiple years) and reformatted them to log 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).<sup>11</sup> 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 of 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, reserves, bodies of water, or other unpopulated areas. 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.<sup>12</sup>

---

<sup>10</sup> [www.cdf.go.ke](http://www.cdf.go.ke), accessed May 28, 2012. Information for six constituencies was missing from the data posted on the CDF Board website.

<sup>11</sup> 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 the administrative location within which the actual project is 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.

<sup>12</sup> See Appendix X.

(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, which took place less than a year before the launch of the CDF program (and can thus be taken to be exogenous to any effects that the program might have subsequently had on election outcomes). 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.<sup>13</sup> 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 incentives to put projects where many people live—either because that will allow the projects to benefit the most people 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.<sup>14</sup>

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 well being or because she calculates that the political returns to a project will be greater if people are

---

<sup>13</sup> 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 this is the case: about 95% of voters do in fact register at the polling station closest to their home.

<sup>14</sup> 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 as that previously discussed variable.



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.<sup>15</sup> 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<sup>16</sup>—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 constituency-level 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). Although novel in political science, such models are commonly used in epidemiology and ecology, where researchers have identified occurrences of disease or locations of particular animal or plant species and seek to understand the

---

<sup>15</sup> For evidence linking ethnicity to patterns of patronage distribution in Kenya, see Burgess et al (2013) and Kramon and Posner (2014).

<sup>16</sup> Forty-three percent of constituencies in Kenya have a level of ethnic diversity less than 0.3 (calculations from data described in Harris (2013)).

distribution of these phenomena in terms of a set of explanatory variables.<sup>17</sup> 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 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 for the incumbent MP in the last election) and the number of observed points (projects). This can be done either conditional on covariates or without them.<sup>18</sup> 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. As noted, the more common practice in studies of distributive politics—and not just in Africa—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

---

<sup>17</sup> For an application of the technique to a political science problem—the spatial location of polluters—see Monogan et al (2013).

<sup>18</sup> We estimate all models using the spatstat package in R (Baddeley and Turner 2005).

we analyze countrywide. 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.

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 in different constituencies, and even at different points in time. 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).<sup>19</sup> 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.<sup>20</sup>

(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 largest number of 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) paved 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 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.<sup>21</sup> 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

---

<sup>19</sup> In all estimates, we standardize the measure for cross-constituency comparability by subtracting the mean and dividing by the standard deviation.

<sup>20</sup> We drop population density in the model that generates the boxplot in the final column because it is mechanically correlated with the number of votes won by the MP. The mean correlation across the 204 constituencies between population density and votes won by the incumbent is .633.

<sup>21</sup> Note, however, that our measure of poverty captures only the rate of poverty in a location, not its depth. 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.

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.<sup>22</sup>

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 many Kenyan electoral constituencies are ethnically homogeneous, in which case 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.<sup>23</sup> These findings suggest that Kenyan MPs do in fact favor their supporters in the allocation of CDF projects.

### *Robustness Tests*

While these findings are compelling, one might wonder whether systematic bias in our measurement of project locations might somehow be responsible for the findings we report. Alternatively, one might wonder if the results were somehow a product of the (somewhat opaque) point process model we employ. To address these concerns, we test whether our findings are robust to an analysis on the subset of CDF projects that are located at schools. School-related projects (which comprise over 50 percent of the projects in our data) have the dual advantage of being located in places about whose geo-locations we can be extremely confident and also being amenable to a straightforward statistical analysis. Since we know the locations of the universe of schools in which CDF projects might have been placed, we can simply code schools in terms of whether or not they received a CDF project and then examine this outcome as a function of the estimated number of votes won by the MP in the immediate neighborhood of the school in a logistic regression.

Table 3 presents the correlations and associated p-values between the point estimates on the “# of votes won by the MP” coefficient in the bivariate point process model with the

---

<sup>22</sup> If poverty rates are higher in rural areas, then the result on our poverty measure is also consistent with the finding, reported above, that projects are less likely to be placed in areas with lower population density. 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.

<sup>23</sup> In generating these estimates, we take the log of the number of votes won by the MP prior to standardizing the measure.

analogous coefficient in the logistic regression. We present the results for seven different subsets of our data, broken down by the share of projects in each constituency that are school-related. The higher the share, the more confident we can be in our ability to compare the point process model results (based on the full set of CDF projects) to the logistic regression results (based on the schools-only sub-sample), but the fewer the number of constituencies in which we can make the comparison. Thus, for example, the first column of Table 3 reports the correlation between the two sets of coefficients (logit and point process model) for constituencies that retained 60 percent of all projects. For those 43 constituencies, the correlation is 0.24 with a p-value of 0.12. As we move down the rows, we include more constituencies in the estimation of the correlation coefficient, though those constituencies discard more projects. Across the range of specifications, we find a positive, and in some cases significant, correlation between the coefficients, suggesting that model choice and focusing on high quality data produce similar cross-constituency results to the preferred model.

(Table 3 Here)

To confirm that our point process model results also hold using only the school-related projects—a subset of projects, recall, in which we can largely rule out placement errors or geographic inaccuracies—we examine a bivariate point process model of project locations on electoral support in 2002 including only CDF projects related to schools. We find a 0.57 correlation between the two sets of coefficients (p-value < 0.001). This higher correlation makes sense when we consider the nature of the comparison a point process model allows, and helps us better understand the attenuation in the correlation between the logistic regression results and the point process model results above. The outcomes in the point process models are the point locations of projects—“presence-only data”—which are contrasted with regions in the constituency where no projects are placed. The logistic regression results truncate the potential locations of “non-allocations” to only those schools reported in the schools dataset. Schools themselves are likely already located in areas where individuals are clustered—villages and towns. As noted earlier, population density is highly correlated with the number of registered voters and, in most cases, with the number of supporters. By focusing only on the question of which schools get CDF allocations versus which do not in the logistic regression, we effectively reduce variation in the independent variable of interest since we exclude areas which may have high support rates but relatively speaking low numbers of supporters due to low population density.

### *Variation*

These general patterns summarized above in the discussion of Figure 2 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 set of constituencies and different way in another set of constituencies (and not at all in yet another set of constituencies)?”

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

(Figure 3 Here)

### **Under What Conditions do MPs Favor Their Supporters?**

The factors that lead MPs to be more or less likely to target their supporters with CDF projects can be divided into four categories. The first relate to the characteristics of the incumbent MP him or herself. There is evidence from research in India 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 the 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 central government’s targeting of) such resources, they may be less dependent on CDF funds and freer 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 may also be relevant. We might therefore expect to find the association between votes and projects to be increasing in the MP's prior political experience, which we proxy through the number of elections the MP had won prior to 2002.<sup>24</sup>

The second set of possibly relevant factors relate to the characteristics of the constituency in which the MP was elected. One is whether the constituency is urban or rural—a distinction we can proxy for through (the log of) population density. This may matter for voters' expectations about the kinds of benefits they will receive from politicians, and hence the allocation strategies that politicians pursue (Nathan 2014).

A second potentially important characteristic is the constituency's degree of ethnic heterogeneity. As noted earlier, whereas some Kenyan electoral constituencies are highly heterogeneous, many are comprised almost entirely of members of a single ethnic community. Insofar as the impetus to reward one's supporters stems from ethnic-based norms of reciprocity, and insofar as these norms are stronger when members of one group live amongst members of another, we might expect to find a stronger association between project placement and political support in more ethnically heterogeneous settings.

Alternatively, the constituency's ethnic heterogeneity might matter because it provides an indication of the likelihood that voters who did not support the incumbent in the last election could be convinced to do so in the future given the proper inducements. In the Kenyan context, where support for candidates from other ethnic groups is infrequent (Gibson and Long 2009), it is plausible that non-coethnics are not likely to be convertible into supporters 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. In such a setting, 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 implication is that the association between votes in the last election and project placement should be weaker in homogeneous than in heterogeneous constituencies.

The third set of factors relate to the characteristics of the contest that brought the MP to power. If the MP won the election by a large margin, and if past supporters can be convinced to support the MP again if they are given CDF projects, then targeting past supporters may be a dependable strategy for winning re-election. We would therefore expect a high margin of victory to be associated with a strong correlation between past

---

<sup>24</sup> We only examine elections between 1992 and 2002. Since there were just two elections during that period (1992 and 1997), the measure of prior political experience therefore takes values of 0, 1, and 2.

support levels and project placement. The level of turnout may also matter insofar as it captures the extent to which the population in the constituency is mobilized politically. When they are, 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 to be positively correlated with a stronger association between votes and project placements.

We test these expectations in the first four columns of Table 4. The first three columns present each category of explanatory factors separately; the fourth pools them in a single model. We find no support for any of the individual-level factors: neither the MP's gender nor her membership in the ruling coalition, nor her political experience seems to matter for the extent to which she targets her supporters with CDF projects.<sup>25</sup> Regarding the constituency-level factors, we find no evidence in these models that the urban or rural nature of the constituency (proxied by the log of population density) matters. But we do find evidence, in keeping with the logic described above, that MPs are more likely to target their supporters in more ethnically heterogeneous settings. Contrary to expectations, however, we find no support in these models for the importance of either the MP's margin of victory in 2002 or the turnout rate.

### *The Spatial Distribution of Political Support*

The fact that our data permit us to estimate the association between political support and project placement in 204 separate constituencies makes it possible to investigate the conditions under which Kenyan MPs target their supporters. The fact that our data also allow us to estimate the level of political support (and opposition) at every point in each constituency makes it possible to do something even more novel: to supplement the explanatory variables we have just been discussing with measures that capture the spatial distribution of political support. We investigate two different spatial measures.

The first, which we term "clustering," is a Gini coefficient that captures how evenly the MP's supporters are distributed around the constituency. Calculated using our raster-level estimates of the number of voters who supported the MP in 2002, the measure takes a hypothetical value of one when all of the supporters are located in a single raster and a hypothetical value of zero when each raster in the constituency contains an equal number of supporters.<sup>26</sup> The clustering of the MP's support matters because it affects the extent to which the MP can feasibly target them with projects. As the measure approaches zero, it becomes impossible to distinguish areas that contain supporters from areas that do not,

---

<sup>25</sup> One plausible reason for our null findings with respect to membership in the ruling coalition is that the nature of the ruling coalition changed markedly midway through our period of study when the National Alliance Rainbow Coalition (NARC) split apart. This caused many sitting MPs who had previously been affiliated with the ruling party to become members of the opposition—effectively changing their coding on this variable. We plan to alter our coding scheme to better capture this change—and, indeed, to exploit it to generate and test observable implications about changes in patterns of partisan targeting—in future iterations of this paper.

<sup>26</sup> Since this measure would be strongly affected by rasters that contain no voters, we exclude all rasters that contain no supporters or opponents.



and, absent the ability to put a project at every point in the constituency, the association between political support and project placement will mechanically go to zero. As the measure approaches one, however, the MP's supporters are bunched in a smaller part of the constituency and it becomes straightforward to target them.

Whereas clustering is about the how the MP's supporters are distributed in space, the second measure, which we term "segregation," is about the spatial relationship between the MP's supporters and his opponents. 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  $m$ .

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

This measure is theoretically important because it captures the extent to which projects targeted at the MP's supporters will also benefit her opponents. All else equal, we would expect an MP to be less willing to locate CDF projects in an area that contains large numbers of voters that did not support her in the last election, even if that area also contains large numbers of voters that did (although we shall return momentarily to a situation in which targeting areas that contain both supporters and opponents might be advantageous). It follows that if a constituency is fairly integrated, and thus contains both areas in which the MP's supporters are living alongside her opponents and areas in which they are living by themselves, we would expect the MP to allocate her CDF funding disproportionately to the latter areas. This will reduce the overall association between past electoral support and project placement—at least compared to a constituency in which her supporters and opponents were completely segregated and she could reward her supporters without simultaneously rewarding her opponents.

The distinction between these two measures is depicted in Figure 4. In the top left of the Figure, the MP's political support is both clustered and segregated. Hence targeting is straightforward (just put all the projects on the left side of the circle) and there is little chance of spillover to non-supporters. In such a setting we would expect to find a strong association between support for the MP and project placement. In the bottom right of the Figure, by contrast, the MP's political support is neither clustered nor segregated. Targeting is quite difficult in such a situation, and the likelihood that projects put in the vicinity of the MP's supporters will also benefit his opponents is high. We would therefore expect a much weaker relationship between past political support and getting a CDF project.

(Figure 4 Here)

The bottom left case (clustered but not segregated) is particularly interesting. Here, targeting is feasible but spillover is likely. Ordinarily, this might dissuade the MP from energetically targeting his supporters (lest his opponents also benefit). But in a setting in which opponents are potentially convertible into supporters—for example, when the constituency is ethnically homogeneous and the MP's supporters and opponents are thus both coethnics—we might expect the incentives for MPs to target their supporters to be especially strong. The implication is that we might expect to find a negative interaction

between segregation and ethnic heterogeneity, signifying 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 constituencies, MPs are more apt to reward supporters even when spatially integrated with non-supporters.

We test these various expectations in columns 5-8 of Table 4. As expected, both the clustering and segregation measures (added to the fully specified model from column 4 in columns 5 and 6 respectively) are strongly positively associated with the tendency for MPs to target their supporters with CDF projects. But something else quite extraordinary happens in column 5. First of all, the explanatory power of the model, as captured in the  $R^2$ , increases more than six-fold, suggesting that the clustering of the MP's supporters constitutes a powerful explanatory factor. Second, some of the other non-spatial variables whose estimated effects had been zero in previous specifications now take on large, statistically significant coefficients. For example, we now see a strong positive relationship between the log of population density and the likelihood that MPs target their supporters. This suggests that, once we account for clustering (which is presumably lower in more densely populated constituencies), partisan targeting is more likely in urban or peri-urban areas. The margin of victory and turnout rate are also now significantly associated with the tendency of MPs to target their supporters. These were both hypothesized relationships, but they were absent until we accounted for the spatial distribution of the MP's supporters.

The introduction of the segregation measure in column 6 has a similar effect. While its impact on the  $R^2$  is more modest (merely a doubling), the margin of victory point estimate jumps in magnitude and the ethnic heterogeneity measure reacquires its significant positive coefficient from the earlier specifications. When the clustering and segregation measures are both included in the same model (in column 7) their effect sizes remain large and highly significant. The only slight reduction in the sizes of their respective coefficients suggests that they are capturing different phenomena. Column 8 includes the interaction between ethnic heterogeneity and segregation. Its negative sign is in keeping with the theoretical expectation, but the point estimate is small and not statistically distinguishable from zero.

Taken together, however, these results demonstrate the salience of spatial variables in accounting for when and why politicians reward their supporters. Not only do the clustering of the MP's political support and the segregation of his supporters and opponents matter in their own right, but the addition of these variables in the analysis soaks up variation that had concealed underlying relationships between other important determinants of partisan targeting. The data requirements to generate these spatial measures are large, so their inclusion will not always be possible. But the implication of our results is that the spatial distribution of political support should be incorporated into

analyses of political allocations wherever possible and that researchers who are not able to include such measures should be mindful that their findings might well be different if they had.

## **Conclusion**

TO BE WRITTEN...

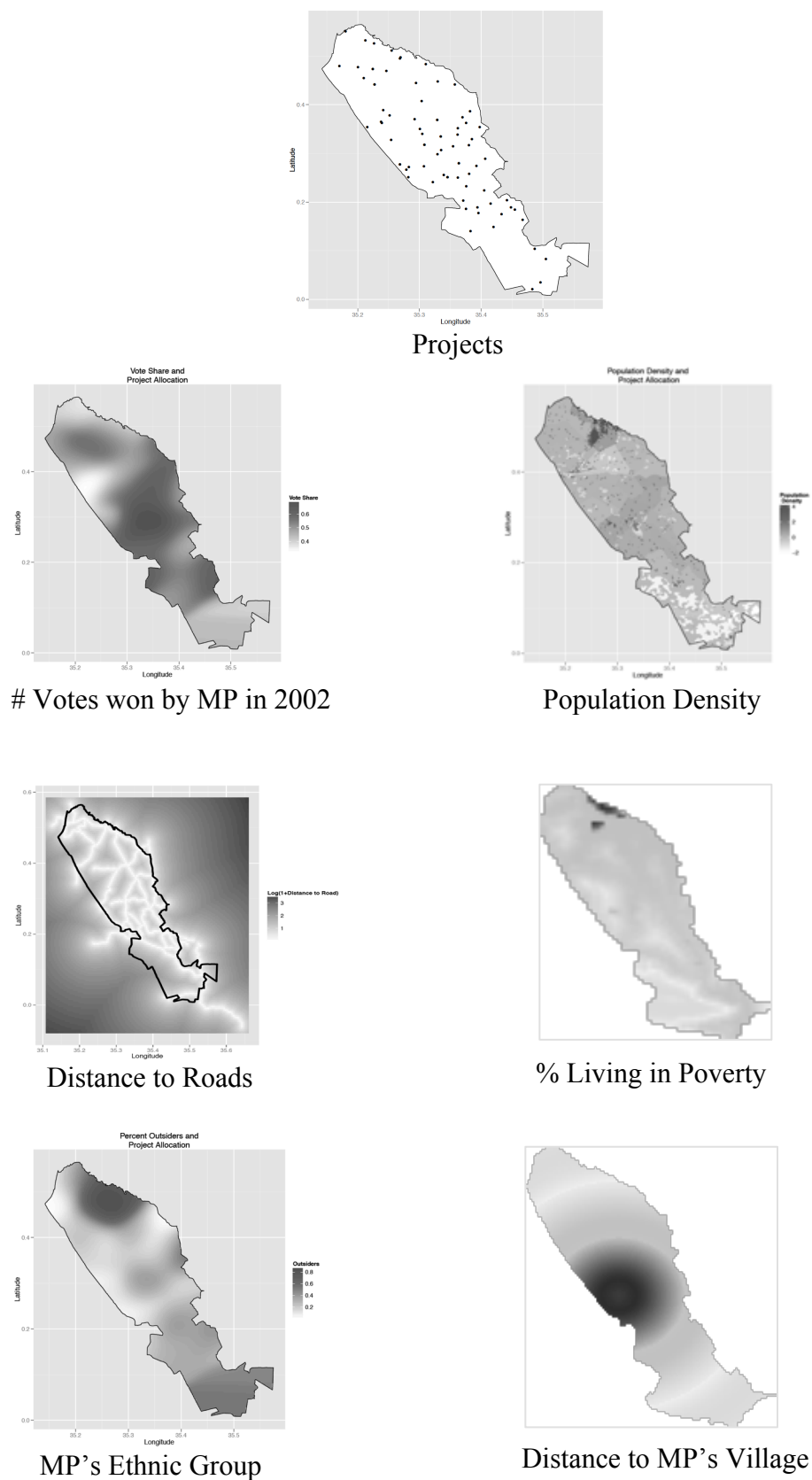
## References

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

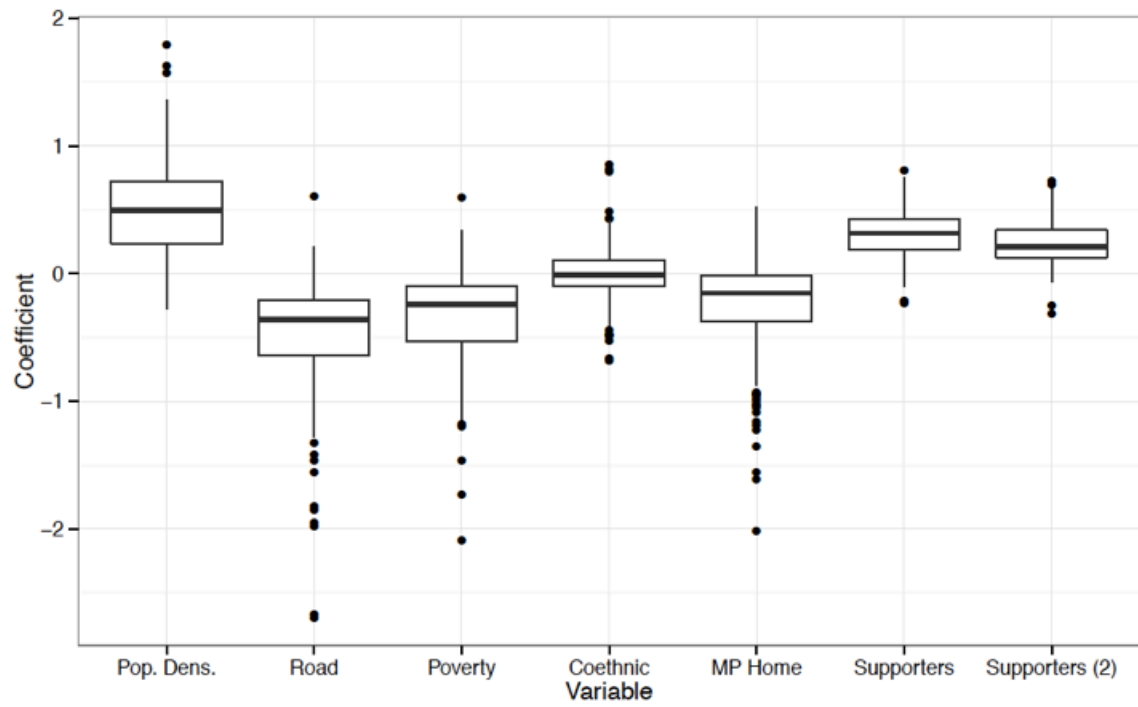
- Franck, Raphael, and Ilia Rainer. 2012. "Does the Leader's Ethnicity Matter? Ethnic Favoritism, Education and Health in Sub-Saharan Africa." *American Political Science Review* 106(2): 294-325.
- Gibson, Clark C. and James D. Long. 2009. "The Presidential and Parliamentary Elections in Kenya, December 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.
- Hoffman, Vivian, Pamela Jakiela, Michael Kremer and Ryan Sheely. 2014. "Targeting, Discretionary Funding, and the Provision of Local Public Goods." 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.
- Keefer, Philip and Stuti Khemani. 2009. "When Do Legislators Pass on Pork? The Role of Political Parties in Determining Legislator Effort." *American Political Science Review* 103(1):99-112.
- 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.
- Monogan, James E. III, David M Konisky, and Neal D. Woods. 2013. "Strategic Placement of Air Polluters: An Application of Point Pattern Models." Unpublished paper.
- Morjaria, Ameet. 2013. "Electoral Competition and Deforestation." Unpublished paper.

- Mutoro, Stephen. 2005. "CDF Has Become a Milch Cow for MPs." *Daily Nation*, 30 March.
- Nathan, Noah. 2014. "Local Ethnic Geography, Expectations of Favoritism, and Voting in Urban Ghana." Unpublished paper.
- 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.
- 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**

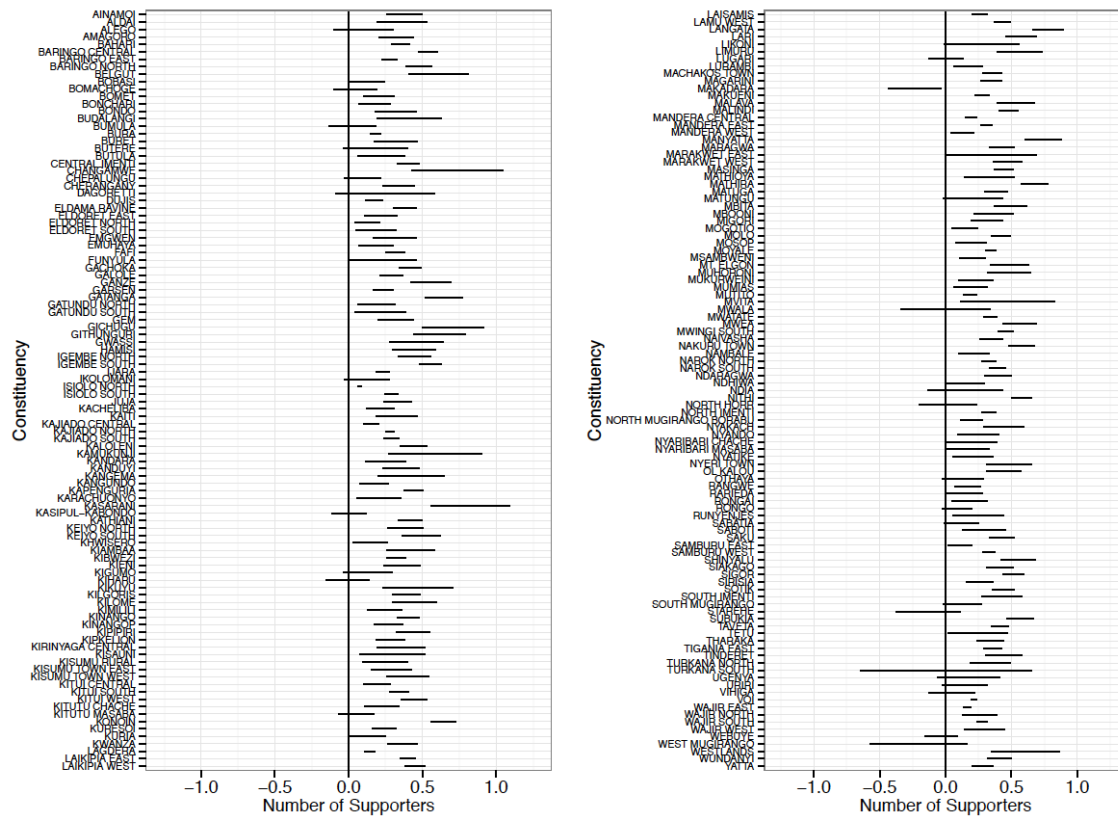


**Figure 2: What Factors Affect Where CDF Projects Are Placed?**

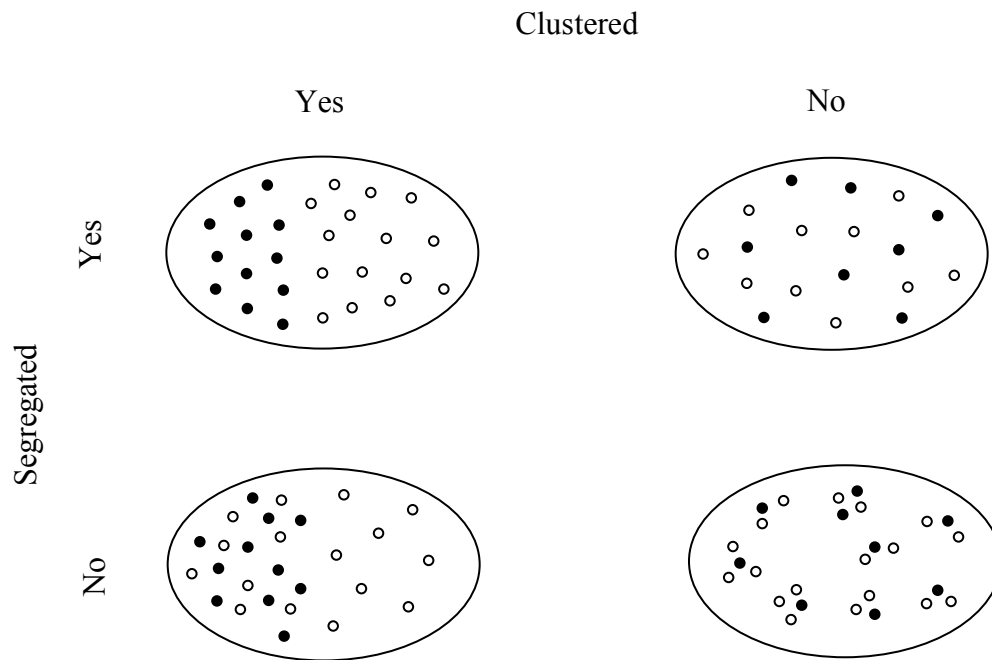




**Figure 3: Association Between Votes Won by the MP and Project Placement, by Constituency**



**Figure 4: Two Dimensions of the Spatial Distribution of Political Support**



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

---

<i>CDF program funding</i>	
Total allocations <sup>1</sup>	US\$ 369,494,222
Avg allocation per constituency	US\$ 351,900 (min=\$66,667; max=\$675,401)
Avg % of allocation spent	77
<i>Projects</i>	
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
<i>Project types (%)</i>	
Health	9
School	54
Water	13
Other	15
<i>Project duration (%)</i>	
1 year	57
2 years	25
3 years	12
4 years	5
5 years	1

---

<sup>1</sup>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: Correlation of Overall Point Process Model Results  
with Schools-only Logistic Regression Results**

% of projects that are school-related	Number of observations	Correlation	p-value
0.6	43	0.24	0.12
0.5	85	0.27	0.01
0.4	132	0.12	0.17
0.3	169	0.09	0.26
0.2	193	0.17	0.02
0.1	196	0.17	0.02
0	202	0.19	0.01

**Table 4: Under What Conditions do MPs Favor their Supporters?**

	<i>Dependent variable:</i>							
	Association between # votes won by MP and project placement							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Characteristics of the MP</i>								
female	0.040 (0.073)			0.008 (0.074)	0.006 (0.067)	0.025 (0.074)	0.020 (0.067)	0.017 (0.067)
member of ruling coalition	0.033 (0.025)			0.035 (0.028)	0.035 (0.026)	0.033 (0.028)	0.033 (0.025)	0.033 (0.025)
# of elections won 1992-2002	0.007 (0.025)			0.004 (0.025)	0.002 (0.023)	-0.004 (0.025)	-0.004 (0.023)	-0.004 (0.023)
<i>Characteristics of the constituency</i>								
log population density		0.005 (0.007)		0.002 (0.008)	0.034*** (0.009)	-0.003 (0.008)	0.029*** (0.009)	0.028*** (0.009)
ethnic heterogeneity		0.076* (0.043)		0.093** (0.046)	0.038 (0.042)	0.097** (0.045)	0.042 (0.042)	0.101 (0.113)
<i>Characteristics of the 2002 election</i>								
margin of victory in 2002			0.041 (0.049)	0.024 (0.052)	0.081* (0.047)	0.279*** (0.114)	0.298*** (0.103)	0.292*** (0.104)
% turnout in 2002			0.040 (0.137)	0.169 (0.150)	0.511*** (0.145)	-0.067 (0.176)	0.300* (0.168)	0.311* (0.170)
<i>Spatial characteristics</i>								
clustering of MP's support					0.701*** (0.104)		0.687*** (0.103)	0.675*** (0.106)
segregation of MP's support						0.901** (0.362)	0.775** (0.328)	0.847** (0.353)
segregation of MP's support * ethnic heterogeneity								-0.296 (0.530)
constant	0.288*** (0.021)	0.257*** (0.042)	0.273*** (0.083)	0.137 (0.110)	-0.662*** (0.155)	0.045 (0.115)	-0.725*** (0.156)	-0.732*** (0.156)
observations	204	204	204	204	204	204	204	204
R <sup>2</sup>	0.011	0.018	0.004	0.034	0.215	0.064	0.237	0.239
adjusted R <sup>2</sup>	-0.004	0.008	-0.006	-0.001	0.183	0.025	0.202	0.199

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

## Appendix

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

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

For  $W$ , we can estimate the intensity as the count of points in  $\mathbf{x}$  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:

$$X \sim \text{Poisson}(\lambda(u))$$

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).