We leverage innovative spatial modeling techniques and data on the precise geo-locations of more than 34,000 Constituency Development Fund (CDF) projects in Kenya to test whether MPs reward their supporters. We find only weak evidence that MPs channel projects disproportionately to areas inhabited by their political allies, once we control for other factors that affect where projects are placed such as population density, poverty rates, ethnic demographics, and distance to paved roads. Notwithstanding this result, we find evidence for significant cross-constituency variation in political targeting, driven in large part by the spatial segregation of the MP’s supporters and opponents. Our findings thus challenge the conventional wisdom about the centrality of clientelistic transfers in Africa and underscore how local conditions generate particular incentives and opportunities for the strategic allocation of political goods. Although based on substantive materials from an African case, the paper thus speaks to fundamental issues in distributive politics.
Political scientists have long debated the strategies that politicians pursue as they allocate the resources they control. Is their objective to reward their supporters (Cox and McCubbins 1986)? Is it to induce people who did not support them in the past (or who did not vote at all in the last election) to support them in the future (Lindbeck and Weibull 1987; Nichter 2008)? Or is their goal simply to help the people with the greatest needs? Adjudicating among these different strategies is challenging because of the difficulty in identifying precisely who is targeted by the transfers and because politicians rarely have full discretion over resource distribution. In addition, most studies focus on the allocation decisions of a single actor (the president, the ruling party, “the government”), which makes it difficult to disentangle systematic from idiosyncratic factors and impossible to investigate the conditions under which particular strategies are more or less likely to be pursued. In this paper, we leverage unique data from a policy innovation in Kenya, the Constituencies Development Fund (CDF), which permits us circumvent these challenges and provides a clean and unambiguous test of the first of these hypothesis: that politicians reward their supporters.

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 resources, we can use the spatial allocation of such projects to test whether their placement is shaped by where their supporters are located. We explore these allocation decisions using a unique dataset on the precise geo-locations of more than 34,000 CDF projects initiated between 2003 and 2007, when the then-incumbent MPs ran for re-election. 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 as well as
highly disaggregated data on population density, the location of roads, poverty rates, and ethnic demographics. We then use spatial modeling techniques to examine the placement of projects within each constituency as a function of these independent variables, generating estimates of the relationship between project location and levels of political support in 196 different constituencies.\textsuperscript{1} 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 varies with the local conditions they face.

We find evidence that Kenyan MPs do not, in fact, favor their supporters in the distribution of CDF projects as strongly or as universally as the literature would lead us to expect. This is because MPs consider other factors as well, such as the number of people living in the area in question, its distance from a paved road, the local poverty rate, the proportion of the population in the area that shares their ethnic background, and proximity to their own village. We show that once we control for these other determinants of project placement the effect of political support diminishes. Equally important, we find that the tendency for MPs to channel CDF projects to their supporters varies from constituency to constituency depending, in particular, on whether or not the MP is politically affiliated with the ruling political coalition, the MP’s margin of victory in the last election, and whether the MP’s supporters and non-supporters are spatially segregated from one another (which, in turn, is affected by the spatial distribution of the population more generally). These findings, based on significantly richer data than has

\textsuperscript{1} While Kenya had 210 constituencies during the period under study, fourteen constituencies were excluded from analysis due to lack of data on either CDF projects or the results of the preceding elections.
been employed in prior studies, challenge the conventional wisdom that African politicians are motivated primarily by a clientelistic bargain in which political support is exchanged for public resources (Chabal and Daloz 1999; Lindberg 2010).

Beyond this important empirical finding, the paper makes three contributions. First, by studying the allocation decisions of constituency-level rather than national-level actors, we can exploit cross-unit variation in the associations we find between project placement and political support. This allows us to go beyond the question of whether politicians favor their supporters to identify the conditions that make such behavior more or less likely. This represents a major advance over prior work in which a single (or collective) decision-maker, subject to a fixed set of conditions, makes the allocation decision. Our analysis, by contrast, explores the allocation decisions of 196 Kenyan Members of Parliament, each operating under the same set of rules and with similar resources, but in a different social, demographic, and political environments.²

Second, the paper joins a handful of recent projects (Hoffmann et al 2015; Chhibber and Jensenius 2016; Bussell 2016; Carlitz forthcoming) that study the allocation decisions of

² Studies that exploit changes in regime type or political leaders—for example, Franck and Rainer (2012), Jablonski (2014), Hodler and Raschky (2014), Burgess et al (2015), and Kramon and Posner (2016)—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. The most similar study to ours is Keefer and Khemani (2009), who study the spending decisions of 483 MPs in India in an analogous program. However, their outcome of interest is the proportion of the allocated money that is spent rather than the spatial distribution of that spending across the MP’s constituency.
politicians at the sub-constituency level using project-level data. Most analyses of
distributive politics focus on the allocation of resources across administrative units such
as provinces (Calvo and Murillo 2004), states (Khemani 2007), districts (Weinstein 2011;
Burgess et al 2015), constituencies (Jablonski 2014), or municipalities (Dahlberg and
Johansson 2002; Weitz-Shapiro 2014; Diaz Cayeros, Estevez and Magaloni 2016).3 In
generating their estimates of favoritism, these analyses implicitly assume that all the
inhabitants of the unit have equal access to the resources that the unit receives. This
assumption ignores the second-order distributive problem that arises once the resources
arrive within the unit. When the resources are in the form of specific projects, it also
ignores the fact that where a project is built 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 or roads in
a district/constituency/municipality may not tell us very much about “who gets what.” By
contrast, by geo-referencing every CDF project, our approach allows us to directly
examine the spatial allocation of development resources. This dramatically minimizes the
ecological inference problem inherent in unit-level approaches and permits more precise
judgments about who benefits from the allocation of resources. Our sub-constituency,
project-level approach also has the benefit of aligning theory—which focuses on within-
district, rather than cross-district, allocations—with empirical tests (Dixit and Londregan
1996; Cox and McCubbins 1986; Cox 2009).

3 Studies that employ DHS data, such as Franck and Rainer (2012) or Kramon and Posner
(2016), 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.
Finally, our geo-coded project data and high resolution information on voting patterns and demographics make it possible to study how the spatial concentration of the people living in a voting district and the segregation of the incumbent’s supporters and opponents 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. Indeed, our findings regarding the importance of population concentration and the segregation of supporters and opponents suggest that studies that fail to account for these spatial characteristics may generate misleading conclusions about how distributive politics operates.4

The Constituencies Development Fund (CDF) Program in Kenya

4 Our work thus joins a growing body of research emphasizing the role of geography in politics. However, unlike this extant literature, which generally focuses on how the distribution of political support affects public policy outcomes (Beramendi 2012; Rickard 2012; Jurado and Leon forthcoming) or how votes translate into seats in different types of electoral systems (Chen and Rodden 2013; Calvo and Rodden 2015; Jusko 2015), our study focuses on the impact of political geography on the distribution of more local public goods. Furthermore, whereas the extant literature emphasizes the distribution of partisan preferences across districts, our analysis examines the distribution of political support within them. The closest work to ours is Ejdemyr, Kramon and Robinson (2015), which finds that MPs in Malawi provide more local public goods when ethnic groups are geographically segregated.
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 poverty rate. Once the funds arrive in each constituency, they are disbursed by a constituency-level CDF Committee, which, during the period we study, was effectively controlled by the MP.\(^5\) Citizens and organized groups are invited (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. As a consequence, CDF funds are widely regarded as *pesa ya Mheshimiwa*, Swahili for “the MP’s money.”\(^6\)

CDF transfers represent a considerable infusion of funds for local development. Total CDF allocations between 2003 and 2007 totaled nearly $333 million (see Table 1). This amounts to an average of $316,709 per constituency per year and, during the period

---

\(^5\) The MP’s control was diminished in 2013 with the passage of an amended Constituencies Development Fund Act. However, the period we study was prior to the tightening of these rules. See section 23 of the Constituencies Development Fund Act of 2003. Further details of the CDF program’s structure and operation are provided in Ndii (2014).

we study, represented roughly 18 percent of the total average development expenditure per constituency on education, roads and public works, health, and water.\textsuperscript{7} Countrywide, CDF allocations funded a total of 34,102 projects over the five-year period we study, or an average of about 162 projects per constituency.\textsuperscript{8}

(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” (GoK 2003) This requirement means that they are spent on local public goods: the rehabilitation of school classrooms, bridge repairs, road grading, water projects, or the construction of local infrastructure such as 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

\textsuperscript{7} Calculated based on approved development expenditure data for 2006/2007 as summarized in Institute of Economic Affairs (2007).

\textsuperscript{8} As indicated in Table 1, an average of 84 percent of the funds allocated were actually spent on projects. Although this implies that some MPs were “passing on pork” (Keefer and Khemani 2009), this figure compares favorably with similar decentralized fiscal transfer schemes in other settings. For example, Keefer and Khemani (2009) report that legislators in India spend 82 percent of the Member of Parliament Local Area Development Scheme funds they were allocated. 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.
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\(^9\) sum of money that, subject to minimal restrictions, may be allocated within the constituency with nearly total discretion.\(^{10}\) This provides the MP with an important tool for rewarding his supporters. Although CDF funds constitute just one component of total local development expenditures, the close identification of CDF with the MP makes it an extremely valuable political resource. The CDF program thus provides researchers with a nearly ideal opportunity to observe how political actors distribute the resources they control. And since we can observe such distribution decisions in nearly 200 separate “laboratories,” we can also draw important lessons about the conditions under which they engage more or less strongly in political favoritism.

\(^9\) A caveat regarding the exogenous nature of the size of the disbursements: although the formula that determines each constituency’s allocation would appear to be tamperproof, it does generate incentives for MPs to lobby the Kenya National Bureau of Statistics to revise their constituency’s official poverty rate so as be able to claim a larger share of CDF funding. Although some MPs have indeed attempted this strategy (Wahome 2008), such lobbying is rarely successful and, when it is it, can only result in a relatively small change in a constituency’s allocation given that poverty rates affect just 25 percent of the total allotment.

\(^{10}\) This stands in marked contrast to the situation faced by members of the U.S. Congress, whose allocation decisions are constrained by the preferences of their parties and the president, and for whom most resources are not discretionary. See Dynes and Huber (2015) for a useful recent review.
As a local public good, the benefit of a CDF project is inversely correlated with a person’s distance from it. This means that we can infer which people an MP is targeting based on where in his constituency he chooses to locate each CDF project.\textsuperscript{11} 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.\textsuperscript{12} And we can do this while controlling for a host of other factors that offer competing explanations for how MPs allocate their CDF funds.

Three potential objections might be raised to this approach. First, one might wonder if it is reasonable to trust the project records we use in our analyses (described below). 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 examples of ghost projects, unaccounted for money, and general mismanagement of CDF program funds during this period (e.g., Mutoro 2005; Lumwamu and Munene 2006; Awiti 2008; National Anti-Corruption Campaign Steering Committee 2008; Ndii 2014). Although we acknowledge this concern, the project lists from which we culled information on CDF projects were compiled and, to the best of our knowledge, verified by CDF managers.

\textsuperscript{11} 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, and projects like bridges constitute only a small share of our sample.

\textsuperscript{12} 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.
hired in 2008 in each constituency. These managers were beholden not to the local MP, but rather to the central CDF oversight board. Thus, short of verifying each project in the field, these lists represent the best available data on actual projects implemented during the period of study.

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 the same location: one can be a refurbished school classroom; another a clinic; another a borehole. 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 (or 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 a completely different area, and even outside the

---

13 Section 21 of the Constituencies Development Fund (Amendment) Act of 2007 establishes these managerial officers and their independence from the MP. In addition, the Act charges the manager with maintenance of all CDF related records in the constituency.

14 Indeed, most of the projects in our data set are placed “on top of” projects that were already constructed in earlier periods: you cannot build a girls’ latrine at a school without there having first been a school.
constituency altogether. Since our inference about whom the MP is targeting depends on the 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) 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 unlikely, given that an MP has many possible opportunities to use a given contract to satisfy both voters and contractors.

**Our Data**

*CDF Project Locations*

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 public 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) and information about the amount of money allocated to the project for the year in question. We coded the sector in which the project was located based on its name. Some reports also include a description of the project’s implementation status (e.g., ongoing, complete, not started). Our principal data, a geo-referenced set of 34,102 CDF projects in 196
constituencies for the 2003 to 2007 financial years, are extracted from this source.\textsuperscript{15} We downloaded the documents—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 the names of facilities for which point or polygon data were available (e.g., schools, market towns, health centers, water/irrigation features).\textsuperscript{16} As indicated in the first row of Table 2, we were able to match 60 percent of projects to their exact geo-referenced locations. In another 21 percent of cases we were able to match the project to its correct census enumeration area (of which there are an average of 32 per constituency, each with a median area of 1.34 square kilometers). In a further 10 percent of cases we were able to

\textsuperscript{15} Accessed May 28, 2012 from www.cdf.go.ke. Information for six constituencies was missing from the data posted on the CDF Board website.

\textsuperscript{16} 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.
match the project either to its administrative sub-location (of which there are an average of about 20 per constituency) or location (of which there are an average of about 12 per constituency). In only 9 percent of cases were we unable to match the project to any point or administrative sub-unit within the constituency. As shown in Table 2, our procedures allowed us to place 69.9 percent of all projects within 1 km of their true location and 75.1 percent within 2.5 kms of their true location—well inside the radius within which residents would benefit from most projects.

(Table 2 Here)

In the cases in which we were not able to match a project to a specific point, we randomly placed the project at a point within the smallest unit to which we could assign it (i.e., its enumeration area, sub-location, location, or the constituency as a whole) with the probability of placing the project at each point proportional to its estimated population density. In effect, this step prevents us from inadvertently locating projects in national forests, parks, reserves, bodies of water, or other unpopulated areas. To account for measurement error in imputed project locations, we created 21 separate data sets of project locations and ran all of the analyses in which project locations are the dependent variable on each of these 21 separate data sets. The results we report below are the average coefficient estimates of these 21 separate regressions, with standard errors calculated following the procedures discussed in King et al (2001).

*Explanatory Variables*

To estimate the spatial association between project placement and political support for the MP, we combine the CDF project data just described with polling station-level electoral returns from Kenya’s 2002 parliamentary elections. These elections took
place a year before the launch of the CDF program and can thus be taken as exogenous to any effects that the program might have subsequently had on election outcomes. We link these results to a geo-referenced polling station dataset to create rasters that identify the estimated number of votes won by the winning candidate at each point in each constituency. We do this by assigning each 100 square meter grid cell in each constituency to its most proximate polling station using a Voronoi diagram. Then we interpolate the level of political support in each grid square by multiplying the share of votes won by the incumbent at the polling station by the grid square’s estimated population, based on the high-resolution population density rasters described in Linard et al (2012). With approximately 14,000 polling stations countrywide and an average of

Formally, we do this by defining each constituency as a region $R$, defined on a $J \times K$ raster reporting the (estimated) number of people in grid square $p_{jk}$. The spatial distribution of political support is then defined by segmenting $R$ into Voronoi diagram polygons corresponding to the areas closest to each polling station, where polling stations are indexed $m = 1, \ldots, M$. $R_m$ refers to the area of region $R$ defined by the Voronoi diagram polygon corresponding to polling station $m$. Next, we generate the spatial distribution of support in two steps. First, within $R_m$ we calculate the percentage of the population in $R_m$ in each grid cell. Then, we multiply the number of supporters of the incumbent at polling station $m$ times the percentage of population in each grid cell in $R_m$ to get $s_{jk}$, which reports the expected number of supporters in grid cell $jk$.

Effectively, this reifies the assumption that, local to the polling station, support is uniformly distributed across the spatial distribution of the population. An alternative assumption would be that support is uniformly distributed across $R_m$. However, this approach is problematic because it does not account for the fact that many areas contain
roughly 67 polling stations per constituency, this procedure generates an extremely fine-grained constituency-level map of the distribution of electoral support for the incumbent MP.\(^{18}\)

In addition to this main explanatory variable of interest, we include several important control variables in our analyses. The first is population density measured at the grid square-level. Population density is important because MPs will have incentives to put projects where large numbers of 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. It also turns out that population density is mechanically correlated with our other explanatory variables. As we describe below, this requires that we adopt a set of estimation procedures that account for this fact.

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.\(^{19}\) Distance to a paved road may be important few to no people, and would therefore cause us to impute the presence of support where it could not possibly exist (e.g., in the middle of a forest reserve or lake).

\(^{18}\) One reasonable concern is that citizens may register to vote in a place that is distant from their place of residence. If that were the case, then a citizen’s electoral support would be detached from the benefits accruing due to local public goods. IED (2002) presents evidence that suggests that this is not a concern. That study asked citizens why they registered at a particular polling place and 94.3% responded that “it was because of the proximity to their place of residence.”

\(^{19}\) Although one might be concerned that these data could include roads that were paved with CDF funds, we think this is unlikely. First, CDF funds tend to be used for grading
insofar as projects located closer to roads are cheaper to build, and, all else equal, we would expect MPs to favor locations that allow them to do the same thing more inexpensively. Since the effect of the distance to the road is not likely to be linear, we use the square of the distance in our analyses.

Although we argued earlier that the need for development projects is great throughout Kenya (and therefore that the allocation of a project to a particular location at time \( t \) would not necessarily reduce the demand for a second project in that same location at time \( t + 1 \)), 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 the MP calculates that the political returns will be greater for projects that are sited where 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 (2014) to create constituency-level rasters that depict estimates of the number of people living in poverty in each one square kilometer grid cell. These data also, of course, put us in a position to test the hypothesis that poverty alleviation, not political strategizing, guides MPs’ decisions about project placement.

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 relationship with voters will already be built into the distribution of his political support. However, it may nonetheless be of interest to test whether projects are disproportionately channeled to the MP’s coethnics. There is ample evidence of such ethnic targeting in dirt roads rather than paving new roads of the sort captured in GoK (2006). Second, roads are at most only four percent of CDF projects.
Kenya at the province- (Barkan and Chege 1989), district- (Burgess et al 2015), and constituency-levels (Jablonski 2014). But, for lack of fine-grained data, there is no evidence for ethnic targeting within constituencies.\footnote{Ejdemyr, Kramon and Robinson (2015) do provide evidence for ethnic targeting at the sub-constituency-level in Malawi.} Part of the reason is also that Kenyan constituencies are often ethnically homogeneous, or nearly so.\footnote{Forty-three percent of constituencies in Kenya have a level of ethnic diversity less than 0.3 (calculations from data described in Harris (2015)).} Also, prior studies focused on ethnic targeting by national political actors rather than locally elected representatives. To test for ethnic targeting by MPs at the constituency level, we employ polling station-level estimates of ethnic demographics from Harris (2015) to create a raster identifying the estimated number of people at each point in the constituency that are members of the MP’s ethnic group.

We also take a second cut at the expectation that MPs will favor members of their own kin group by creating a raster for each constituency indicating the square of the distance from each point to the MP’s home village.\footnote{Technically, we calculate the square of the distance from the polling station at which the MP is registered to vote rather than the square of the distance to his MP’s home village. But these are in almost every case the same thing, and it is more intuitive to think of the measure in terms of the distance to the MP’s home village.} Especially in ethnically homogeneous constituencies, the relevant communal distinction may not be between coethnics and non-coethnics (since there are none, or very few, 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 their ancestral home, allocating development projects there will raise their future comfort level.

The distance-to-the-MP’s-home-village measure also serves a second purpose. The models we use to estimate the extent to which MPs favor their supporters take advantage of extremely fine-grained information about the spatial distribution of each MP’s political support in the last election. While it is reasonable to think that MPs have a rough sense of what this distribution looks like, their perceptions may be inaccurate—or at any rate not as accurate as the data we are using to model their behavior. This could lead to a disconnect between our assumptions about the information they are using to guide their placement of CDF projects and the actual information that affects the spatial allocation decisions we observe. However, even where an MP has only a hazy understanding of how his support is distributed, he can still make a strong inference that his support is high in the area surrounding his home village. The distance-to-the-MP’s-home-village measure therefore provides the basis for a low-information complement to our main analysis that utilizes the fine-grained information on the distribution of the MP’s supporters.

Figure 1 provides an illustration of what our data look like for an example case:

---

23 The polling station-level election returns we utilize are publicly available, so any MP who wanted to access the data we employ could do so. However, we think it is unlikely that any but a handful of MPs have taken advantage of these data to map distribution of their political support.

24 He would be right to do this: our data suggest that, on average, MPs received 77 percent of the vote in the polling station closest to their home village (compared to an average of 62 percent in other polling stations).
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.

(Figure 1 Here)

**Modeling the Association Between Project Locations, Political Support, 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 1996; Diggle 2003). Such models are commonly used in fields such as epidemiology, ecology, geology, and meteorology where researchers have identified occurrences of a phenomenon of interest (for example, outbreaks of a disease or locations of particular animal or plant species, mineral deposits, or weather patterns) and seek to understand the spatial distribution of these phenomena in terms of a set of explanatory variables. However, we know of only two applications of the technique to political science problems: Warren (2015)’s study of radio and cell phone penetration and collective violence in Africa and Monogan et al (In Press)’s study of the spatial location of air polluting firms in the United States.25

The intuition behind how the point process model works is straightforward. Begin by projecting onto a two-dimensional geographic space an interpolated gradient of some

---

25 Monogan et al (In Press) model the locations of air polluters relative to hazardous waste polluters. Given the lack of relevant control observations analogous to hazardous waste polluters in the Monogan et al (In Press) study, we model the intensity of projects as a function of underlying spatial covariates.
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 and complement these points with an arbitrary number of “dummy” points representing locations where projects are not located. Then use the actual and dummy points to divide the space into polygons that contain outcome information about both the location of dummy and project locations in each polygon, as well as spatial covariate information of the sort described in the previous section. 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), either conditional on covariates or without them.26 Strictly speaking, we treat the distribution of CDF projects as an inhomogenous Poisson process—one that varies across space as a function of the social, spatial, and political attributes described above.

The point process approach has two distinct advantages over standard techniques for studying political allocation decisions. First, as noted, it allows us to model spatial variation in project placement directly without relying on ad hoc aggregations such as districts, constituencies, villages, or other administrative units. In Kenya, as in many other places, the borders of such units are products of political processes (Oucho 2002; Kasara 2006), so using those borders to delineate units of analysis could introduce endogeneity or aggregation bias. Second, the point process approach 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. This represents a significant advance over the common practice of aggregating outcomes and

26 We estimate all models using the spatstat package in R (Baddeley, Rubak and Turner 2015). For further technical details, see Appendix A.
explanatory variables to larger units or imposing an arbitrarily sized grid over the area studied and coding the outcome of interest as present or absent in each grid square, since both such approaches obscure the more localized relationship between project placement and characteristics that may vary (sometimes quite considerably) within these units. While the severity of this ecological inference problem will depend on the sizes of the units to which the explanatory and outcome variables are aggregated—with analyses done at the grid square- or village-level presenting much less of a problem than analyses done at the level of the constituency or district—it will always be preferable to avoid the problem altogether by estimating the association, as we do here, at the level of the project units themselves (Amrhein 1995).

*Population Density*

One complexity that arises from the move to a point-level analysis is that CDF projects are very unlikely to be placed where there are no people. This presents a problem because, as Figure 2 demonstrates, the vast majority of the Kenya’s land area is uninhabited (or very nearly so), and this highly skewed distribution of people generates a strong mechanical correlation between population density and our explanatory variables, including, crucially, the number of supporters at a given point in space. Although it is in principle possible for two places with the same number of people living in them to have different levels of support for the incumbent (one could be the MP’s own village; the other an opposition stronghold), the fact that the overwhelming majority of places have no people (and hence no supporters) means that any estimate of the association between supporters and project placement will be driven largely by population density. The same is true for the number of people living in poverty, the number of people belonging to the
MP’s ethnic group, and the distances from roads and the MP’s home village.

(Figure 2 Here)

These strong correlations are clearly demonstrated in Figure 3, which plots the grid square-level correlation in each constituency between the log of population density and each of our main explanatory variables, with statistically significant correlations plotted as black dots. Given these strong relationships, it becomes impossible to separate out the effect of each variable from the effect of population density.

(Figure 3 Here)

To deal with this problem, we regress (the log of) population density on each of our explanatory variables and then use the residuals from these regressions in lieu of our direct measures of each variable. This allows us to interpret the estimated spatial association between each explanatory variable and project placement as capturing the effect of the part of each variable that is not due to population density.

**What Factors Affect Where CDF Projects are Placed?**

The fact that our data enable us to estimate the association between political support and project placement in 196 separate constituencies presents a challenge for how to present our findings. One approach would be to estimate a pooled model across all 196 units and report the results of that single analysis. However, 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 model. This is at odds with both the evidence we present below and our priors, which are that MPs face different local conditions and will thus pursue different strategies of project allocation in different constituencies.
We therefore eschew a pooled analysis and instead present the distribution of
county-level estimates for each variable of interest. Each boxplot in Figure 4
presents the estimated association between project placement and the variable in question
for each of the 196 constituencies in our analysis. Constituencies in which the estimated
relationship is significantly different from zero are plotted as black points. As
noted, the estimates are the average of 21 separate regressions, each using a slightly
different set of interpolated project locations.

(Figure 4 Here)

Population density: The first column plots the association between population
density and project placement. As expected, there is a statistically significant positive
association in nearly all constituencies. As noted, this is in large part mechanical: very
few projects get placed in areas with few or no people. But it also stems from that fact
that, where there are people, MPs seem to allocate CDF projects to areas where they are
present in the largest numbers. In the median constituency (shown by the dark horizontal
bar in the center of the boxplot), a one standard deviation increase in the log of
population density is associated with an additional 0.25 projects allocated to that area.
This pattern is consistent with the interpretation that MPs are seeking either to benefit
the largest number of people or to win the most votes. Combined with the strong
correlations between population density and the other explanatory variables shown in
Figure 3, the clear relationship between population density and project placement justifies

27 We standardize the estimates for cross-constituency comparability by subtracting the
mean and dividing by the standard deviation.
our strategy of using the residuals in lieu of the direct measures of the variables we turn
to next.

*Distance to paved roads:* The association between project placement and distance
to 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. Our
estimates suggest that, for the median constituency, a one standard deviation increase in
the distance from a road squared leads to a decrease in intensity of projects placed in that
area by approximately 0.24 projects.

*Poverty:* The relationship between poverty and project placement is more mixed.
Although there are some constituencies in which the association is positive (and a
handful in which the relationship is statistically significant and negative), in most
constituencies CDF projects are no more likely to be placed in areas with larger numbers
of poor people. Thus, contrary to the CDF program’s stated objectives (and to at least one
leading hypothesis for why politicians allocate resources as they do), CDF funds do not
appear to have been used as a tool for poverty alleviation. ²⁸

Another interpretation of these results, however, 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

²⁸ Note that our measure of poverty captures only the rate of poverty in a location, not its
depth. We therefore cannot completely rule out that MPs are channeling projects to the
very poor—providing that they live in isolation from larger concentrations of less poor
people.
do, apply for projects). Consistent with this interpretation, Baird et al (2013) provide evidence from a similar program in Tanzania that poverty is negatively associated with the likelihood of applying for project funding. The lack of a statistically significant association between poverty and project placement in many constituencies may simply reflect the offsetting effects of these two processes.

*Ethnicity:* In keeping with the received wisdom about the central role of ethnicity in Kenyan politics, the vast majority of MPs locate CDF projects in areas dominated by their coethnics. Our estimates suggest that, in the median constituency, a one standard deviation increase in the number of coethnics in an area translates to an approximate increase of 0.27 projects (in log counts). This finding is all the more striking given that many Kenyan electoral constituencies are ethnically homogeneous, which makes targeting projects along ethnic lines quite difficult.

*Distance to the MP’s village:* We also find evidence that MPs in many (but, somewhat surprisingly, not all) constituencies are more likely to locate CDF projects in proximity to (i.e., at smaller distance from) their home village. As noted, in addition to providing support for the idea that politicians allocate the resources they control to favor their closest kin, this finding also offers evidence for our main outcome of interest, partisan favoritism, if we believe that MPs have little information about where their strongest supporters are located but infer (correctly) that their support is likely higher in their home villages than elsewhere.

*Political support in the last election:* We turn now to the association between project placement and the number of votes won by the MP in the last election. The bivariate relationship is significant and positive in the vast majority of constituencies, suggesting that MPs are indeed more likely to put CDF projects in areas where their
supporters are. In the median constituency, a one standard deviation increase in the number of supporters translates into a 25% increase in the likelihood of project placement over an area of average support, all else held constant. However, this relationship weakens significantly when we control for the other factors that we have just shown are also associated with project placement: the square of the distance to paved roads, the number of people living in poverty, the number of the MP’s coethnics, and square of the distance to the MP’s home village (and, of course, indirectly, population density). Once we have accounted for the effect of these other factors there is no relationship in the median constituency between the level of political support the MP received in the last election in a given area and the likelihood that he places a CDF project there. As noted, this finding runs strongly against the conventional wisdom that politicians in Africa can be expected to use the resources they control to favor their own.

However, this null effect in the median constituency masks significant cross-constituency variation. In places like Yatta in Machakos district, Kacheliba in West Pokot district, Kieni in Nyeri district, and Butula in Busia 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 Kathiani in Machakos district, Wundanyi in Taita Taveta district, and Mbita in Suba district (to select but a few of the constituencies with zero coefficients) we find no evidence that MPs were any more likely to channel CDF projects to areas in which more people voted for them in the last election. In still other constituencies (for example, Malindi in Malindi district or Ndia in Kirinyaga district), we estimate a statistically significant negative relationship between political support and project placement, suggesting that MPs were more likely to put CDF projects in places where their support was weaker. These findings suggest that the
interesting question may not be whether politicians reward their supporters but under what conditions.

Under What Conditions do MPs Favor Their Supporters?

The factors that might lead MPs to be more or less likely to target their supporters with CDF projects can be divided into three categories.

Characteristics of the MP

The first set of factors relates to the personal characteristics of the incumbent MP. 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 ways. For opposition MPs, on the other hand, CDF funds constitute the lion’s share of the development resources they personally control—indeed, the CDF program itself was proposed by the
political opposition (Ndii 2014). Hence we might expect to find a stronger association between votes won and project placement in constituencies where the MP is not a member of the ruling coalition.\footnote{Coding membership in the ruling coalition is complicated by the fact that many MPs switched parties midway through our study period following the defeat of a government-supported constitutional reform proposal in a national referendum in 2005. Our coding rule is to assign an MP to the ruling coalition if they were allied with the governing party either before or after the party system shake-up. This implies that our estimates should be taken as a lower bound on the effect of membership in the ruling coalition.}

Insofar as knowing how to deploy resources to maximize one’s re-election prospects requires political experience, an MP’s status as a neophyte versus an 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 via a dummy variable that takes a value of 1 if the MP was an incumbent when he ran in 2002. Alternatively, incumbent MPs may have already channeled public goods to their supporters during their prior term (although not via the CDF program, which did not exist until 2003), so the need for additional projects in their areas of strongest support may be lower. This would imply a negative relationship between the number of prior terms served and the association between political support and project placement.

\textit{Characteristics of the Electoral Contest that Brought the MP to Power}

The second set of factors relate to the characteristics of the electoral contest that brought the MP to power. The first of these is the closeness of the election. If the MP won by a small margin, then he is likely to feel vulnerable and to be extremely strategic...
in how he deploys his CDF funds. To the extent that he believes that favoring his supporters will increase his chances of re-election, this should generate a strong association between past support and project placement. If the margin of victory is comfortable, by contrast, he may feel he can win re-election without focusing overly closely on rewarding his past supporters with CDF spending, leading to a weak relationship between past levels of support and project placement. Indeed, as the margin of victory increases beyond a certain point, the distinction between supporters and opponents evaporates (since everyone is a supporter) and it becomes mechanically impossible to discriminate between the two in the allocation of project support. For this reason, we operationalize margin of victory as a dummy variable capturing whether the margin of victory in 2002 was above or below the median.

Another crucial factor (which we are able to examine only because of the unique, fine-grained spatial data we are employing) is the degree to which the MP’s supporters and opponents are geographically segregated from one another. The segregation of supporters and opponents matters because CDF projects are public goods that benefit everyone in the vicinity in which they are located. If we define the targeting of supporters as providing more benefits to supporters than opponents, then targeting is only possible when opponents and supporters live in different places. If they are living right on top of one another, it becomes impossible to locate projects that benefit one without also benefiting the other.³⁰

If we define targeting of supporters somewhat differently, as providing benefits to supporters irrespective of whether these benefits are also enjoyed by opponents, then the

³⁰ Ichino and Nathan (2013) explore the implications of this point for ethnic voting in Ghana.
spatial relationship between supporters and opponents might seem not to matter: the MP can simply put projects where his voters are and ignore whether constituents who did not vote for him may also benefit. However, from the perspective of the researcher (or the voter trying to infer whether the MP is making good on his promise to reward his supporters), spatial proximity matters quite a lot. This is because when supporters and opponents are living in the same place, favoritism and the absence of favoritism (or even a strategy of rewarding of opponents) become observationally equivalent. It is only possible to infer whether an MP is targeting his supporters when those supporters are segregated from non-supporters and the MP can feasibly distribute projects to one and not the other.

The implication is that, irrespective of how we define targeting, it is only possible when supporters and opponents are segregated. When they are living in close proximity, targeting is either mechanically impossible (if we adopt the “rewarding supporters without also rewarding opponents” definition) or impossible to infer from the observed pattern of resource distribution (if we adopt the “rewarding supporters irrespective of whether non-supporters might also benefit” definition). Either way, our expectation is that the constituency-level relationship between votes and project placement should be positively associated with segregation.\(^\text{31}\) To measure segregation, we adopt the spatial

\(^{31}\) A second expectation (which follows from the greater indeterminacy in constituencies with low segregation) is that we should find a stronger relationship between segregation the association between votes and project placement if we restrict our sample to constituencies where segregation is high. Notwithstanding the fact that truncating the sample in this way attenuates the relationship, we find that the relationship is indeed stronger in the restricted sample (analyses not shown).
information theory index developed by Reardon and O’Sullivan (2004). For our purposes, the measure takes a value of 1 when an MP’s supporters and opponents are maximally segregated from one another and a value of 0 when they are as mixed as they possibly could be given the total distribution of partisanship.

**Characteristics of the Constituency**

The third category of factors that might affect the extent to which MPs favor their supporters relate to the characteristics of the constituency in which they were elected. One potentially salient 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, a 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 it is rare for voters to support candidates affiliated with parties that are perceived to represent other ethnic groups (Posner 2007; Gibson and Long 2009), 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 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, again, that the association between votes in the last election and project placement should be weaker in homogeneous than in heterogeneous constituencies.

Another salient characteristic is the degree to which the constituency’s population is geographically clustered. As we saw in Figure 2, the vast majority of areas in Kenya are unpopulated. But the ways in which people are spread over space varies considerably from constituency to constituency, and this matters a lot for an MP’s ability to target his supporters—largely because of its mechanical effect on segregation. When a constituency’s population is spread broadly over space, it is possible for supporters to be located in one area and opponents in another. However, when a constituency’s population is highly clustered, supporters and opponents are necessarily close together, so segregation (and the ability to target or to measure targeting) will be low. Hence we expect the association between votes and project placement to be negatively related to population clustering. We measure population clustering in terms of relative entropy, which we operationalize by comparing the grid square-level populations we observe in each constituency with the hypothetical situation in which the population was evenly distributed across all grid squares. High values of our measure imply tight clustering of people across space, whereas low values imply a more even distribution.

Results

We test these various expectations in Table 3, both in specifications without province-level fixed effects (columns 1, 3, 5, and 7) and with them (columns 2, 4, 6, and 8). The dependent variable is the constituency-level association between project
placement and the number of supporters in the last election, which is our measure of political favoritism. The estimates we use are conditional on our other main covariates, as depicted in the last column of Figure 4. We look first at each category of explanatory factors separately, and then in a pooled model in columns 7-9.

(Table 3 Here)

Turning first to the candidate’s own characteristics (rows 1-3), we find that the association between votes and project placement is weaker in constituencies with female MPs. We caution, however, that this finding is driven by a very small number of MPs—there are just six females in our sample—who are all atypical politicians with national-level reputations for whom the regular rules of political behavior may not apply. Hence we hesitate to draw general conclusions from these results about the systematic impact of gender on political favoritism.

We find that membership in the ruling party is associated in all of our specifications with lower levels of targeting of supporters. As suggested earlier, this is likely because members of the ruling party have access to other sources of funding to distribute, and this weakens their need to rely on the strategic allocation of CDF. We find that MPs who had previously served in office are no more likely to channel resources to their supporters—perhaps because of the offsetting effects of greater experience and the lesser need to reward one’s supporters due to distributions already made during the prior term.

Somewhat surprisingly, we find that MPs elected in closer races are no more likely to reward their supporters than those elected in more lopsided contests (although, as we discuss below, margin of victory has important conditional effects). Targeting of supporters is, however, strongly associated with whether or not supporters are spatially
segregated from non-supporters. In keeping with the expectations outlined above, we find that the constituency-level association between votes and project placement is higher where segregation is higher.

Turning next to the characteristics of the constituency (rows 6 and 7), we find no evidence that ethnic heterogeneity is related to the degree of partisan targeting. But we do find that targeting is more likely where people are less spatially clustered. As explained above, this is likely because population clustering is negatively correlated with segregation: only if people are not all living in the same place is it possible for supporters and opponents to be sufficiently far apart for politicians to target one without also targeting (or seeming to target) the other. The fact that the effect of population clustering loses its statistical significance when we add province fixed effects is likely due to province-level variation in spatial characteristics and population density.

Although analytically distinct, segregation and margin of victory interact in important ways. Segregation is a necessary condition for targeting one’s supporters, but absent a small margin of victory the MP may not have incentives to target. Hence we might expect to find a negative sign on the interaction between segregation and margin of victory. This is precisely what we find. The interpretation of the results in column 9 are that the general effect of segregation doubles in constituencies where the margin of victory is very small but goes to zero in constituencies where the incumbent won in a landslide.32

32 One reasonable criticism of our approach here is that the estimates in Table 3 do not account for the fact that the outcome variable is an estimated coefficient with a standard error. Following Lewis and Linzer (2005), we account for this both by using heteroscedastic consistent standard errors and by implementing a weighted least squares
Conclusion

In this paper, we leverage data on the precise locations of more than 34,000 Constituency Development Fund projects in Kenya to explore whether MPs reward their supporters. Using innovative spatial modeling techniques and extensive fine-grained geo-referenced data on election results, population density, poverty rates, ethnic demographics, and the locations of roads, we find only weak evidence that MPs channel projects disproportionately to areas inhabited by their political supporters, once we control for other factors that affect where projects are placed. This finding presents a challenge to the conventional wisdom about the centrality of clientelistic transfers in Africa, and of the use of CDF programs in particular as vehicles for MPs to direct resources to their supporters (Baskin and Mezey 2014). It suggests that some of the partisan patterns we think we see in the distribution of public resources in Africa are in fact driven by underlying correlations between partisanship and other factors.

However, notwithstanding this general result, we find evidence for significant cross-constituency variation within Kenya. A major contribution of the paper is to explain this variation—and, in so doing, to underscore that different local conditions generate different incentives and opportunities for the strategic allocation of political model where the weights are one over the standard error, meaning that more uncertain estimates are down-weighted in the analysis. In both of these additional analyses (included in Appendix B), the signs and magnitudes of the results are unchanged. More importantly, for the most important findings on segregation and population clustering, the results remain statistically significant. However, for the results on the female indicator and the interaction between segregation and vote margin, the results fall above the 10% threshold.
goods. Our most important finding in this respect is the critical role played by electoral geography, which we show provides a key, and largely overlooked, structural constraint on the ability of political actors to play the game of distributive politics. We demonstrate that the spatial segregation of supporters and opponents strongly conditions an MP’s ability to reward the people who voted for him in the last election. This finding suggests that future theory development on distributive politics—both in Africa and in the Political Science discipline more broadly—must consider how politicians’ select strategies based on such fundamental geographic constraints, rather than assuming that distributive strategies involving local public goods are available to all.33

33 Ejdemyr, Kramon and Robinson (2015) take an important first step in this direction by investigating the impact of ethnic segregation on levels of public goods provision and favoritism in their allocation.
References


Baddeley, Adrian. 2010. “Analysing spatial point patterns in R.” CSIRO, University of Western Australia.


Beramendi, Pablo. 2012. *The Political Geography of Inequality: Regions and*


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

<table>
<thead>
<tr>
<th>CDF program funding</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total allocations</td>
<td>US$ 332,544,799</td>
</tr>
<tr>
<td>Average allocation per constituency</td>
<td>US$ 316,709</td>
</tr>
<tr>
<td>Average % of allocation spent</td>
<td>84</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Projects</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of projects</td>
<td>34,102</td>
</tr>
<tr>
<td>Average per constituency</td>
<td>163</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Project types (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>School</td>
<td>64</td>
</tr>
<tr>
<td>Water</td>
<td>12</td>
</tr>
<tr>
<td>Health</td>
<td>8</td>
</tr>
<tr>
<td>Other</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Project duration (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year</td>
<td>56</td>
</tr>
<tr>
<td>2 years</td>
<td>24</td>
</tr>
<tr>
<td>3 years</td>
<td>12</td>
</tr>
<tr>
<td>4 years</td>
<td>6</td>
</tr>
<tr>
<td>5 years</td>
<td>1</td>
</tr>
<tr>
<td>Polygon</td>
<td>Point</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------</td>
</tr>
<tr>
<td>% Matched</td>
<td>60</td>
</tr>
<tr>
<td>Imputation area (sq. km)</td>
<td>1</td>
</tr>
<tr>
<td>% Matched</td>
<td>69.9</td>
</tr>
</tbody>
</table>
## Table 3: Under What Conditions do MPs Favor their Supporters?

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient of project placement and residual support, conditional</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Female</td>
<td>-0.16**</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
</tr>
<tr>
<td>Member of ruling coalition</td>
<td>-0.05**</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
</tr>
<tr>
<td>Incumbent</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
</tr>
<tr>
<td>Margin of victory (2002)</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
</tr>
<tr>
<td>Segregation of supporters</td>
<td>0.14**</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
</tr>
<tr>
<td>Ethnic heterogeneity</td>
<td>-0.20</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
</tr>
<tr>
<td>Population Clustering</td>
<td>-0.04*</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
</tr>
<tr>
<td>Margin of victory x Segregation</td>
<td>-0.27*</td>
</tr>
<tr>
<td>Constant</td>
<td>0.03*</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
</tr>
<tr>
<td>Observations</td>
<td>196</td>
</tr>
<tr>
<td>Province FE</td>
<td>No</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.06</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.04</td>
</tr>
</tbody>
</table>

*Note:* * p<0.01; ** p<0.05; ***p<0.01
Figure 1: An Illustration of our Data in a Sample Constituency

- # Votes won by MP in 2002
- Population Density
- Distance to Roads Sq
- # Living in Poverty
- # of MP’s Coethnics
- Distance to MP’s Village Sq
Figure 2: The Highly Skewed Distribution of Population in Kenya
Figure 3: The Correlation Between Log Population Density and Our Other Variables
Figure 4: What Factors Affect Where CDF Projects Are Placed?
Appendix A: Technical Details of the Point Process Model\textsuperscript{34}

The observed data in our analysis are the locations of \( n \) CDF projects, \( x = \{x_1, \ldots, x_n\} \), whose spatial distribution is a realization of the point process \( X \) in a given constituency \( R \); \( x \in R \). The Poisson process model estimates parameters of the *intensity function* for all locations \( u \in R \). The intensity function is:

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

where \( E[N(X \cap B)] \) is the expected number of points in \( B \), a region within \( R \). For \( R \) we can estimate the intensity as the count of points in \( x \) divided by the area of \( R \). 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.

We define \( \lambda(u) \) is the intensity of a local Poisson process at location \( u \). Note that covariates \( Z \) are measured at every point in \( R \). 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 \( u \). Definition of units of analysis for estimation in this framework follow from the point nature of the data. Two kinds of points are used: project locations and “dummy” point representing spatial locations without any projects. This combined set of points form a *quadrature scheme* that breaks up the area of analysis \( R \) into disjoint spatial units that can be analyzed using familiar Poisson log-linear regression.

\textsuperscript{34} This explanation closely follows Baddeley (2010).
In our analysis, we make two choices regarding the model defaults. Although these choices do not affect the substantive results, we report them here for transparency. First, we face a choice regarding the number of dummy points to include in each constituency-level point process model. A higher of dummy points leads to a more stable estimate, but at not insignificant computational cost. Ideally, we would set the density of dummy points identically for all constituencies. However, this approach would lead to a computationally impractical number of dummy points for large constituencies (e.g., virtually anywhere in North Easter Province). As a result, we vary the number of dummy points used as a flexible function of constituency area. To do so, we calculate the bounding box of the constituency (in meters), and set a quantity $Q$ equal to the longest dimension of that bounding box divided by 100. Then we set the spacing of dummy points equal to $\max(Q, 250)$. This ensures that, for large constituencies, we retain a relatively fine grid of dummy points (ensuring high approximation of two dimensional space). For small urban constituencies, this ensures that the dummy points are spaced 250 meters apart.

Second, several methods exist for estimating the parameters of interest, including maximum pseudolikelihood, logistic likelihood, variational Bayes likelihood, and the Huang-Ogata method. We use the maximum pseudolikelihood method, as it is equivalent to the maximum likelihood in the case of Poisson regression and is unbiased in the presence of a large number of dummy points (like the number we specify). See Baddeley and Turner (2000) and Baddeley and Turner (2005) for details on the software.
### Appendix B

#### Table B1: Table 3 Revisited Using Heteroscedastic Consistent Standard Errors

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient of Project Placement and Residual Support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>−0.16</td>
<td>−0.15</td>
<td>−0.15</td>
<td>−0.16</td>
<td>−0.16</td>
<td>−0.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.15)</td>
<td>(0.16)</td>
<td>(0.13)</td>
<td>(0.15)</td>
<td>(0.16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Member of Ruling Coalition</td>
<td>−0.05**</td>
<td>−0.08***</td>
<td>−0.07**</td>
<td>−0.08***</td>
<td>−0.08***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incumbent</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vote Margin</td>
<td>0.02</td>
<td>0.01</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segregation of Supporters</td>
<td>0.14*</td>
<td>0.16*</td>
<td>0.13*</td>
<td>0.16*</td>
<td>0.29***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnic Heterogeneity</td>
<td>−0.20</td>
<td>0.05</td>
<td>0.08</td>
<td>0.18</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.37)</td>
<td>(0.51)</td>
<td>(0.34)</td>
<td>(0.52)</td>
<td>(0.51)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population Clustering</td>
<td>−0.04*</td>
<td>−0.03</td>
<td>−0.06**</td>
<td>−0.06*</td>
<td>−0.06*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vote Margin × Segregation</td>
<td>0.03*</td>
<td>0.05</td>
<td>−0.01</td>
<td>−0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.04)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.03</td>
<td>0.05</td>
<td>0.01</td>
<td>0.05</td>
<td>0.01</td>
<td>0.08</td>
<td>0.07</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.04)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
</tr>
<tr>
<td>Province FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R²</td>
<td>0.06</td>
<td>0.10</td>
<td>0.02</td>
<td>0.05</td>
<td>0.02</td>
<td>0.04</td>
<td>0.11</td>
<td>0.14</td>
<td>0.15</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.04</td>
<td>0.05</td>
<td>0.01</td>
<td>0.05</td>
<td>0.01</td>
<td>0.08</td>
<td>0.07</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.04)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

#### Table B2: Table 3 Revisited Using Weighted Least Squares

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient of Project Placement and Residual Support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>−0.06</td>
<td>−0.07</td>
<td>−0.07</td>
<td>−0.09</td>
<td>−0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Member of Ruling Coalition</td>
<td>−0.04**</td>
<td>−0.06***</td>
<td>−0.05**</td>
<td>−0.07***</td>
<td>−0.07***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incumbent</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04*</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vote Margin</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Segregation of Supporters</td>
<td>0.10</td>
<td>0.13*</td>
<td>0.13*</td>
<td>0.17***</td>
<td>0.24***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.06)</td>
<td>(0.07)</td>
<td>(0.06)</td>
<td>(0.07)</td>
<td>(0.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnic Heterogeneity</td>
<td>−0.05</td>
<td>−0.04</td>
<td>0.10</td>
<td>−0.03</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.30)</td>
<td>(0.39)</td>
<td>(0.30)</td>
<td>(0.38)</td>
<td>(0.38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population Clustering</td>
<td>−0.02</td>
<td>−0.03</td>
<td>−0.05**</td>
<td>−0.06**</td>
<td>−0.05*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vote Margin × Segregation</td>
<td>0.03*</td>
<td>0.04</td>
<td>0.004</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0.04</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.04)</td>
<td>(0.01)</td>
<td>(0.03)</td>
<td>(0.01)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.04</td>
<td>0.06</td>
<td>0.01</td>
<td>0.03</td>
<td>0.01</td>
<td>0.02</td>
<td>0.08</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td>196</td>
</tr>
<tr>
<td>Province FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R²</td>
<td>0.04</td>
<td>0.06</td>
<td>0.01</td>
<td>0.03</td>
<td>0.01</td>
<td>0.02</td>
<td>0.08</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.02</td>
<td>0.01</td>
<td>0.002</td>
<td>−0.02</td>
<td>−0.002</td>
<td>−0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.04)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td></td>
</tr>
</tbody>
</table>

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