Segregation, Ethnic Favoritism, and the Strategic Targeting of Local Public Goods[1]

SIMON EJDEMYR
Stanford University

ERIC KRAMON
George Washington University

AMANDA LEA ROBINSON
The Ohio State University

November 3, 2015

[1]The authors thank Armstrong Chavula and Bright Chimatiro for valuable research assistance, and Blessings Chisinga, Daniel Young, and the National Statistics Office of Malawi for sharing data and expertise. We also thank Francisco Garfias, David Laitin, Salma Mousa, Adrienne Lebas, Brigitte Zimmerman, and seminar participants at George Washington University, Ohio State University, Stanford University, Texas A&M University, University of Gothenburg, University of Pittsburgh, the 2014 American Political Science Association conference, and the 2014 African Studies Association conference for comments on earlier versions of this paper.

Abstract

This article demonstrates that ethnic segregation is a key determinant of public goods provision. We argue that this results from politicians' strategic engagement in ethnic favoritism: only when ethnic groups are sufficiently segregated can elites efficiently target their coethnics with local public goods. We test this expectation with fine-grained data from Malawi on the spatial distribution of ethnic groups, geolocated distributive goods (public and private), and the ethnic identities of political elites. We find that members of parliament provide more local public goods to their constituencies when ethnic groups are geographically segregated, but that this increased investment is primarily targeted toward coethnics. Thus, while segregation promotes overall public goods provision, it also leads to greater favoritism in the distribution of these goods. Our logic and evidence provide an elite-driven explanation for the considerable variation in ethnic favoritism and the under-provision of public goods in ethnically diverse settings.

Word Count: 9,940 words

The expectation that political elites seek to favor their ethnic kin has long been a staple in the study of African politics (Bates 1983; Joseph 1987). A number of empirical studies have confirmed that coethnics of African political leaders are favored in the distribution of public goods like health care, schooling, and infrastructure (Burgess et al. 2015; Franck and Rainer 2012). Yet there is also considerable variation in whether and to what extent public goods are subject to ethnic favoritism (Franck and Rainer 2012; Kramon and Posner 2013). Why is there ethnic favoritism in the distribution of public goods in some contexts but not in others?

We demonstrate that ethnic segregation helps answer this question. While a substantial literature examines how segregation relates to political participation and political attitudes (Enos 2014; Kasara 2013; Key 1949; Oliver and Wong 2003; Robinson 2015), the literature on ethnicity and public goods provision has generally overlooked the spatial distribution of ethnic groups. Furthermore, this literature has not theorized about how politicians' incentives influence public goods provision, focusing instead on factors that promote or discourage collective action (Alesina, Baqir, and Easterly 1999; Habyarimana et al. 2009; Miguel and Gugerty 2005). We argue that when incumbents have reasons to favor their own group (Bates 1983; Carlson 2015; Ekeh 1975), segregation will promote investments in local public goods and lead to more ethnic favoritism in the distribution of these goods. This is because targeting coethnics with local public goods — which are locally non-excludable but costly to access from distant locations — is difficult unless ethnic groups are spatially segregated.

We use administrative records on the location of all new water wells ("boreholes") built between 1998 and 2008 in Malawi to test these expectations. We focus on the provision of boreholes because Malawian members of parliament (MPs) exert substantial discretion over investments in these goods within their single-member constituencies, though we also evaluate the distribution of health clinics and schools. To determine whether ethnic segregation affects MPs' investment decisions, we construct a constituency-level measure of segregation using census data on the ethnic make-up of more than 12,000 localities, small geographic units (2.5 square miles) nested within constituencies. We also determine for each locality within each constituency whether the MP is ethnically matched with the locality's largest group, as well as the proportion of the population that is coethnic with the MP. With these measures, we evaluate how segregation impacts investments in local public goods *across* constituencies as well as ethnic favoritism in the distribution of these goods *within* constituencies.

After accounting for ethnic diversity, population size, population density, and other potential confounders, we find that ethnic segregation is associated with higher public goods provision across constituencies. The magnitude of our estimates is comparable to the estimated effect of ethnic diversity, among the most robust predictors of public goods provision. Using a difference-in-differences design, we then show that localities which ethnically match with the MP are the primary beneficiaries of increased public goods investments in segregated constituencies. Our estimates suggest that, in segregated constituencies, ethnically matched localities are 15-20 percentage points more likely to receive a new borehole than localities not matched with the MP. We find no evidence of such favoritism in integrated constituencies.

Collective action mechanisms cannot account for these patterns. Given that segregated constituencies consist of many homogeneous local communities, segregation could promote public goods provision because segregation proxies for low levels of local ethnic diversity. However, we show that communities that are ethnically matched with the constituency's political representative benefit more, on average, than communities not matched with the MP. This is inconsistent with a collective action mechanism, which would not predict a difference in homogeneous communities' ability to acquire goods based on their ethnic match with the MP.

We also discuss and rule out other potential concerns. First, we allay concerns that unobserved factors, such as politician quality, are driving our constituency level results by showing that segregation does not affect the provision of private transfers (agricultural subsidies). Second, we show that the results also hold for health clinics and schools, increasing our confidence that our theory generalizes beyond boreholes (Kramon and Posner 2013). These results also suggest that, while segregation's impact on borehole provision levels off at moderate levels of segregation, its effect on clinics and schools — goods with higher geographic reach — is most pronounced at extreme levels of segregation. Third, we provide evidence that residential sorting is unlikely to explain our results given patterns of rural migration within Malawi.

By demonstrating the heretofore unacknowledged importance of segregation in shaping overall investments in public goods, as well as favoritism in their allocation, this paper contributes to research on public goods provision, ethnic politics, and distributive politics. First, we introduce a top-down explanation for the well-established negative relationship between ethnic diversity and public goods provision. While we are careful not to discount bottom-up mechanisms, our results demonstrate that politicians' have fewer incentives to invest in local public goods in diverse, integrated settings. Second, we advance a growing literature linking ethnic demography to political outcomes (Ichino and Nathan 2013; Kasara 2013). Ichino and Nathan (2013) show that members of local ethnic minority groups are likely to vote across ethnic lines in anticipation of benefiting from ethnic favoritism toward the majority group. Our findings are consistent with this expectation, and indeed show that voters are right to anticipate public goods targeting towards coethnics. But we also go further by highlighting the conditions under which voters should expect such favoritism, namely when ethnic groups are spatially segregated. Third, our findings have implications for the distributive politics literature, which emphasizes that political elites often favor some groups over others (Golden and Min 2013). Sometimes these groups are defined ethnically, while in other contexts they follow caste, partisan, or religious cleavages. Regardless of how groups are defined, we show that their spatial distribution can help us understand when politicians will use local public goods to engage in favoritism.

Segregation and Local Public Goods Provision

We build on the distributive politics literature to make predictions about how ethnic segregation impacts local public goods provision.[1] Our theory has four components: elite incentives to favor their own ethnic group, budget constraints, the cost structure of local public goods, and ethnic segregation.

Incentives for Ethnic Favoritism

Three features of the political environment in much of Africa create incentives for ethnic favoritism. We are not, however, claiming that politicians get *no* political returns from allocating goods to non-coethnics.[2] Our theory only requires that the political or personal returns to coethnic provision are higher than the returns to non-coethnic provision.

The first incentive for ethnic favoritism arises from differences in politicians' ability to effectively target groups of voters with material benefits. As Dixit and Londregan (1996, 1134) note, politicians' greater understanding of some voters "translates into greater efficiency in the allocation of particularistic benefits." This relative efficiency defines a "core" constituency (Cox and McCubbins 1986). The theoretical literature highlights that politicians are likely to favor their core constituencies in contexts where ideological or programmatic differences between parties are

[1]"Local public goods" are *locally* non-rivalrous and non-excludable, but costly to access from distant locations.

[2]Voters sometimes support non-coethnic politicians (Ichino and Nathan 2013; Conroy-Krutz 2012) and politicians sometimes allocate local public goods to non-coethnic voters. However, our objective is not to explain all instances of public goods provision. Rather, we show that, in the many contexts in which politicians do have incentives to favor their own group, ethnic segregation will be consequential for public goods provision.

small (Cox and McCubbins 1986; Dixit and Londregan 1996), as is largely the case in Africa (Posner 2005).

In much of Africa, core supporters are ethnically defined: politicians are able to allocate distributive goods more efficiently to coethnics than to non-coethnics.[3] Studies providing evidence consistent with this claim include Carlson (2015), who finds that Ugandan voters disproportionately reward the provision of services by coethnic politicians; Wantchekon (2003), who finds that clientelist appeals are more effective when delivered by coethnics; and Kramon (2013), who finds that vote buying in Kenya is more effective when politicians target coethnics. These results are likely driven by several factors. Strong expectations of ethnic favoritism or distrust of out-group politicians may cause voters to discount or ignore the provision of resources by non-coethnic elites (Bates 1983; Carlson 2015; Posner 2005). Politicians may also be better at engaging politically useful intermediaries in their ethnic home areas (Kasara 2007). Intermediaries can enhance the efficiency of resource distribution by providing elites with greater knowledge of their coethnics' preferences and by monitoring and mobilizing communities to ensure that they support the incumbent (Nichter 2008; Stokes et al. 2013).

Second, there are broader strategic considerations that may drive coethnic targeting. The literature on neo-patrimonial politics highlights that there is often an ethnic calculus to coalition-building (Joseph 1987; van de Walle 2003). African presidents allocate executive cabinet positions to elites from different ethnic groups in exchange for regime support or the delivery of ethnic voting blocs. These posts come with opportunities for rent-seeking and discretion over the distribution of jobs and resources. Since cabinet positions are typically allocated to elites who can deliver the support of their ethnic community, MPs have incentives to maintain strong support amongst their coethnics in order to enhance their pre- and post-election bargaining position (van de Walle 2003).

---

[3] "Efficiency" here refers to the electoral returns received (the output) given some input of time and resources.

Third, there are social and psychological drivers of coethnic favoritism. Political elites often face strong informal pressures to take care of their "own" (Lindberg 2003). Voters generally expect to benefit when their coethnics are in power (Posner 2005), and elites may lose social standing or face social sanctioning if they fail to deliver (Bates 1983). In Ghana, for example, Lindberg (2010, p. 136) finds that "everyday tools of shame, harassment, collective punishment of the family, and loss of prestige and status" all serve as informal pressures on MPs. Moreover, consistent with social identity theory (e.g., Tajfel and Turner 1985), elites may derive psychological benefits from favoring in-group members (Ekeh 1975).

Budget Constraints

Our second component highlights politicians' budget constraints. While incumbents often have incentives to disproportionately serve coethnics, they can choose to do this in a variety of ways. In addition, there are many other political activities they could engage in, such as legislating or raising campaign contributions. Limited time and resources mean they must prioritize some activities over others. Thus, even if they have the discretion to build new local public goods within their constituency, they may not do so if they think other activities carry higher political returns. It is therefore necessary to understand the conditions under which politicians are motivated to allocate local public goods. The final two components of our theory jointly specify such conditions.

Cost Structure of Local Public Goods

The cost structure of local public goods is important for understanding when politicians will be motivated to invest in them. Local public goods have relatively high fixed costs but relatively low marginal costs, compared to private goods, which have low start-up costs but constant marginal costs. For example, in comparison with a cash transfer or agricultural subsidy, the fixed cost of building a local public good, such as a borehole or health clinic, is relatively high, and tends to increase with size of the catchment area of beneficiaries. But once that local public good is constructed, additional beneficiaries come at almost no extra cost. This implies that politicians will prefer to invest in local public goods only when they benefit a sufficient number of (electorally responsive) residents in a given local community (i.e., coethnics).

Ethnic Segregation

In sum, African politicians often have incentives to favor their own ethnic group, must choose among many potential strategies to do so, and will choose local public goods only if these goods will benefit a sufficient number of coethnic residents. In these conditions, ethnic segregation should be highly consequential for local public goods investments.

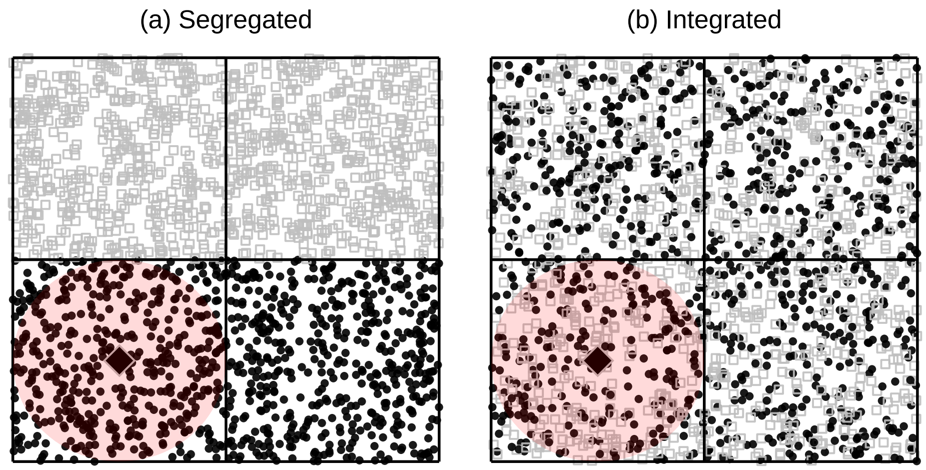
The logic is demonstrated in Figure 1, which shows two hypothetical constituencies with identical levels of diversity, population size, and population density, but different residential patterns of a politician's coethnics (dots) and non-coethnics (squares). As the figure makes clear, local public goods benefit many coethnics when they are geographically segregated, but fewer when coethnics are geographically dispersed. Because many coethnics benefit under segregation, there is a higher chance that the relatively high fixed cost of the good can be justified under segregation than under integration. This logic generates our first observable implication:

H1: *Investments in local public goods will be greater in ethnically segregated constituencies.*

If greater investments in local public goods in segregated constituencies are indeed driven by incentives for ethnic favoritism, local public goods should be disproportionally allocated to coethnic localities within those constituencies. This expectation is consistent with prior research on ethnic favoritism in public goods provision (e.g., Ichino and Nathan 2013; Burgess et al. 2015), but adds the expectation that segregation will shape the degree of ethnic favoritism:

H2: *Within constituencies, ethnic favoritism in the distribution of local public goods will increase with ethnic segregation.*

Figure 1: Two Hypothetical Electoral Constituencies with Different Levels of Segregation



Note: The points and squares represent a politician's coethnics and non-coethnics, respectively, and the black lines separate each constituency into four "localities." The diamond shows the location of a local public good. The transparent circle represents its catchment area. Ethnic diversity, population size, and population density — three predictors of public goods investment — are held constant across the constituencies.

The theory does not necessarily imply a strictly linear relationship between ethnic segregation and the provision and targeting of local public goods. If the reach (catchment area) of a local public good is relatively limited, it may benefit the same number of coethnics at moderate and high levels of segregation. Thus, the impact of segregation on the provision of relatively small-scale public goods, like boreholes or pit latrines, may level out at moderate levels of segregation. With larger-scale public goods such as roads or hospitals, however, we would expect the effect of segregation to keep rising beyond moderate levels. Given their greater reach, these goods benefit more coethnics at high levels of segregation than at moderate levels, making it more likely that politicians can justify their higher fixed costs at high levels of segregation.

Malawian Context

We test our theory using disaggregated data gathered in the ethnically diverse country of Malawi.[4] As in many African countries, Malawi has an institutional structure in which politicians exert significant leverage over the allocation of public goods. We focus on members of Malawi's unicameral parliament (MPs), who are elected by plurality vote in 193 single-member electoral constituencies. MPs play a crucial role in the planning, funding, and management of local public goods in their constituencies. Formal responsibility for the provision of these goods lies with District Assemblies, which by law comprise MPs and locally-elected councilors (Chinsinga 2005). However, local-level elections for councilors were not held until 2000 and after their first term expired in 2005, councilors were never again elected during the period under study. Thus, local development initiatives were largely left to centrally-appointed district officials and to MPs (Chasukwa, Chiweza, and Chikapa-Jamali 2014). MP ability to deliver public goods was heightened in 2006, when parliament introduced constituency development funds (CDF), which provides resources for MPs to engage in development projects in their constituencies.

MPs also exert considerable *informal* influence over the allocation of public goods. As local "big men," they lobby for and influence development projects funded by the central government and NGOs (Cammack et al. 2007; Chasukwa, Chiweza, and Chikapa-Jamali 2014). As a result of this discretion, MPs have increasingly focused on delivering development projects (Chinsinga 2007; 2009), a trend that mirrors dynamics in other parts of Africa (Lindberg 2010). The growing connection between MPs and local development has led to an entrenchment of patronage politics by which MPs exchange tangible goods for political support (Cammack et al. 2007; Chinsinga 2007).

[4] Chewa are the largest group (33%), followed by the Lomwe (18%), Yao (14%), Ngoni (12%), Tumbuka (9%), and smaller groups (Government of Malawi 2008). There is significant variation in segregation across Malawi (Figures SI.1 and SI.2 of Supporting Information).

This pattern of MP patronage is especially true in the provision of new water wells — "boreholes" — for two reasons. First, demand for boreholes is high across Malawi (DeGabriele 2002), with almost half of all rural Malawians having no access to a protected water source in 1998 (Government of Malawi 1998). Second, the relatively low cost of boreholes — approximately \$5000 per good (Baumann and Danert 2008) — means MPs exert substantial influence on where they are built. MPs sometimes do so using CDF or personal funds: in Dowa, for example, an MP was hailed by constituents for drilling 125 boreholes over three years using "personal money through her development office" (*The Malawi Voice* 2014). MPs also impact the placement of government funded borehole projects by, for example, lobbying the Ministry of Irrigation and Water Development or through their formal membership on district councils. Lastly, MPs influence the placement of boreholes provided by other actors, such as NGOs, through informal pressure or partnerships. One Malawian MP described this process as going "shopping for people who can assist" once she could no longer "draw any more money from my pocket because I am dry" (Gilman 2009, p.198). MPs often claim and receive credit for such projects: the MP for Zomba-Likangala was credited with building a borehole despite the funds being provided by an international NGO (*The Nation* 2012).

For both these reasons — the demand among constituents and the ability of MPs to meet that demand — we focus our analysis primarily on the allocation of new boreholes. An additional benefit of focusing primarily on borehole provision is that it allows us to hold constant the cost of providing the public good, as well as the geographic scale of its beneficiaries.

Data and Measurement

We assemble data at two different geographic levels. Our smallest units of observation are 12,380 census enumeration areas, which we call "localities." On average, 1,000 people reside in these localities (Table SI.1 in Supporting Information (SI)). Because the localities are small — on average

2.3 square miles — the catchment area of many local public goods crosses locality boundaries.[5] Our theory thus predicts that the decision to provide a public good to a given locality will depend on that locality's ethnic connection to the political leader *and* on the political leader's ethnic connection to surrounding localities. Our second units of observation are 193 electoral constituencies, within which localities are nested.[6] The average population of constituencies is about 67,000, though a standard deviation of 31,000 indicates substantial malapportionment (Table SI.2).

We construct three key measures. First, we extract the geographic coordinates of all boreholes from the 1998 and the 2008 censuses. We construct a constituency-level count of *new* boreholes built during that ten-year period (on average, 39). We also create a dichotomous locality-level measure that indicates whether each locality received a new borehole from 1998 to 2008 (33%).

Second, we assemble an original dataset on the ethnic identity of each Malawian MP who served between 1994 and 2009 (details in SI). We combine this information with census data on the ethnic make-up of each locality.[7] We create two indicators of the MP's ethnic linkage with a locality. First, we code *Binary Match* equal to 1 if the largest ethnic group within a locality had the same ethnicity as the MP of the constituency at anytime between 1999 and 2008, and 0 otherwise. More than 70 percent of localities were matched at some point. Ethnically matched localities exist in large numbers in constituencies at all levels of ethnic segregation, making it possible for MPs to favor ethnically matched localities in even the least segregated settings. Second, we calculate *Continuous Match* as the share of the locality's population that was coethnic with the MP. If the

---

[5]Ichino and Nathan (2013) estimate that rural Ghanaians can access local public goods 18 miles away.

[6]Constituencies are nested within 28 administrative districts.

[7]For each locality in the 2008 census, we know the total population and the proportion of the population belonging to each ethnic group. While it would be ideal to measure ethnicity prior to 1998, the 1998 census did not ask about ethnicity. We discuss the possibility of residential sorting in the Alternative Explanations section.

ethnicity of the MP changed between 1999 and 2008, we average the two coethnic population shares.[8] On average, 59% of a locality's population is coethnic with the MP (SD=36%).

Finally, we calculate a measure of ethnic segregation for each constituency based on the ethnic demography of all localities within it. We employ the spatial dissimilarity index (Reardon and O'Sullivan 2004), a widely used measure of segregation, which ranges between 0 and 1 with higher values indicating greater segregation.[9] Using this index, we measure how segregated the MP's ethnic group is from other ethnic groups in each constituency across the two legislative terms in 1998-2008.[10] If the ethnicity of the MP changed between the legislative terms, we average across the two MP-specific segregation measures from each term. We do not measure segregation for the ten most ethnically homogenous constituencies ($ELF < 0.05$): ethnic segregation is only theoretically meaningful with at least some ethnic diversity, and a small number of minority group members exert an undue influence on segregation measures in low diversity contexts (Reardon and O'Sullivan 2004).[11]

To illustrate what our segregation measure captures, Figure 2 shows that two constituencies with similar levels of diversity (ELF scores of 0.51 and 0.65) can differ markedly in their degree of segregation (segregation scores of 0.70 and 0.21).[12]

---

[8] Results are robust to analyzing the data separately by term (Table SI.9 in SI).

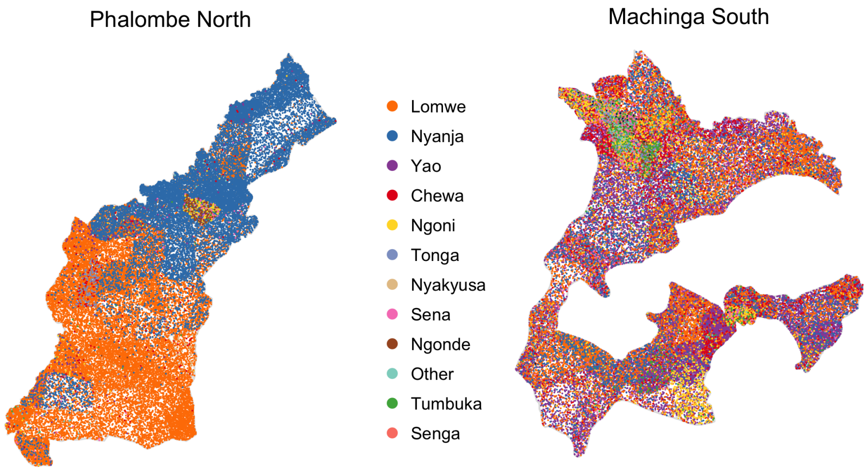
[9] See SI for details.

[10] This MP-specific measure of segregation is more relevant to our theory than an aggregate measure of segregation across all groups. In practice, our measure is highly correlated with a group-size weighted measure of segregation ($r = 0.97$).

[11] Results are robust to including all constituencies (Tables SI.4 and SI.5 in SI).

[12] Our fine-grained measure of ethnic segregation improves upon past research, including Franck and Rainer (2012), who find no evidence that ethnic favoritism by African presidents is conditioned by country-level segregation. Franck and Rainer's segregation measure is based on the geographic mapping of language groups that assumes that language groups have clearly defined boundaries

Figure 2: Ethnic Segregation in Two Constituencies



Note: This figure provides an example of two constituencies with similar levels of diversity but different segregation scores. The spatial dissimilarity score for the MP's ethnic group is 0.70 in Phalombe North and 0.21 in Machinga South. Each dot represents one individual (colored according to ethnicity).

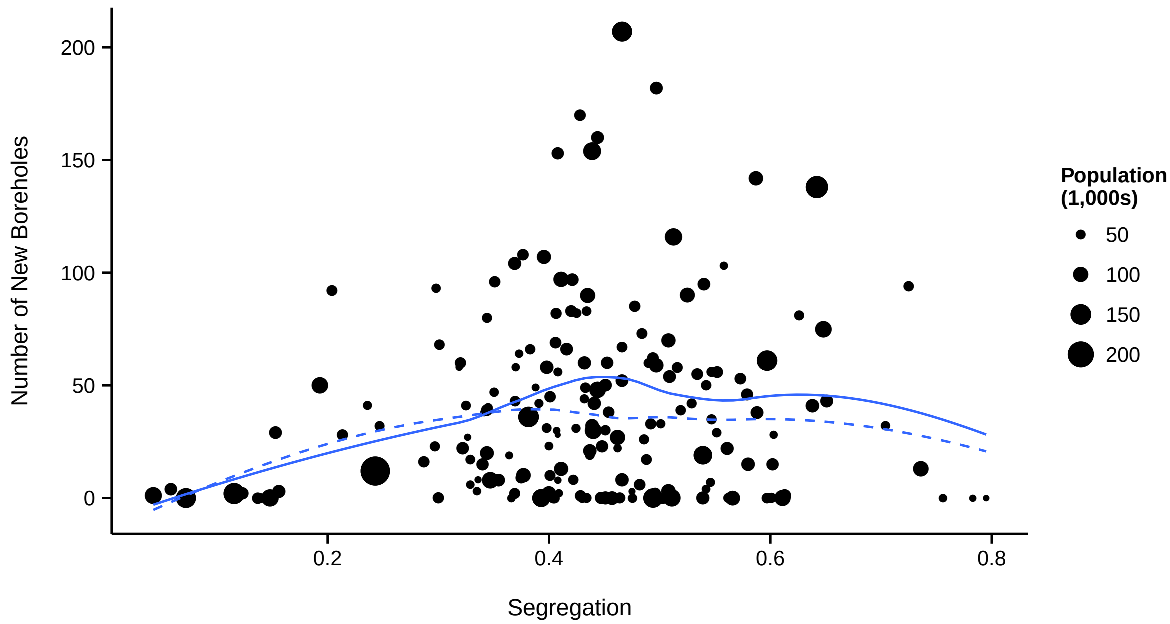
Segregation and Public Goods Provision Across Constituencies

Our theory predicts that investments in local public goods should be higher in ethnically segregated constituencies (**H1**). We first evaluate this expectation with raw data. Figure 3 plots the number of boreholes built in 1998-2008 against ethnic segregation across Malawi's electoral constituencies. Consistent with our hypothesis, new borehole investments and ethnic segregation are positively correlated.[13]

with no overlap (see Matuszeski and Schneider 2006). Thus, by design the data cannot observe ethnic integration that occurs from members of more than one group living in the same local area, an important source of variation in our own data.

[13]The three constituencies with the highest levels of segregation have extremely small populations and did receive new boreholes, pulling the loess curve down at very high levels of segregation.

Figure 3: Segregated Constituencies Invested in More Boreholes than Integrated Constituencies in 1998-2008



Note: This figure shows the relationship between segregation and borehole investments across 183 constituencies in Malawi. Ten constituencies are not shown because they are very homogeneous (ELF scores below 0.05). Point size is proportional to constituencies' population size. The lines are population-weighted loess smoothers; the solid line uses all 183 observations while the dashed line excludes 8 constituencies with very high (2 standard deviations above the mean) borehole investments.

We next use a regression framework to demonstrate that the association between segregation and borehole investment is robust to adjusting for several potential confounders. Because our outcome variable is the count of boreholes built in a constituency, we use a Poisson model modified to account for overdispersion in the data (Gelman and Hill 2006; Wooldridge 1997).[14] We model

[14]This approach allows us to relax the assumption that the conditional variance and mean are equal. While the coefficient estimates from our model mirror those from a standard Poisson model, this approach guards against understating the standard errors.

the number of new boreholes a constituency receives (y_c) as follows::

$$y_c \sim \text{overdispersed Poisson}\left(\theta_c,\ \omega\right), \quad \theta_c = \exp\left(\alpha_{d[c]} + \beta Seg_c + X_c'\gamma\right) \tag{1}$$

where ω is an overdispersion parameter estimated from the data, and where c indexes electoral constituencies and d administrative districts. Our main variable of interest is *Seg*, which measures ethnic segregation. In equation (1), *Seg* is continuous, which assumes a linear relationship (on a log-count scale) between segregation and borehole investments. To allow for non-linearities, we also present results from a model that includes two dummy variables indicating medium and high segregation, leaving constituencies with low segregation as the omitted reference category. We code these categories based on the terciles of the spatial dissimilarity index: low segregation is below 0.401, medium between 0.401 and 0.490, and high above 0.490. We include in vector X_c a set of constituency-level covariates that are likely predictors of borehole investments, including ethnic diversity, population size (in 10,000s), population density (in 1000s per square kilometer), and the number of boreholes in 1998. Since population density is a proxy for urbanization — and demand for boreholes is likely to be less in urban areas — the latter two controls help account for differences in the demand for boreholes across constituencies. We also include administrative district fixed effects, $\alpha_{d[c]}$, because important decisions, including borehole allocation, are often made at the district level.[15]

The results in Table 1 show that segregation is a robust positive predictor of new borehole investments. The coefficients on segregation are positive, statistically significant, and substantively large. Given Model 2, and holding covariates at their mean or mode, we expect constituencies with medium segregation to invest in 11 more boreholes than constituencies with low segregation,

[15]Unlike most fixed effects generalized linear models, the Poisson fixed effects model is a case in which the parameters that maximize the unconditional log likelihood are numerically identical to those that maximize the conditional log likelihood (Greene 2005).

with a 95% confidence interval (CI) of 3 to 22.[16] Similarly, constituencies with high segregation are expected to receive 9 more boreholes than low segregation constituencies (95% CI: 1-22). The effects are comparable to the effect of ethnic diversity: highly diverse constituencies (80^{th} percentile on ELF) invest in 13 fewer boreholes, on average, than low diversity constituencies (20^{th} percentile), with a 95% CI of -33 to -2.[17]

Segregation and Ethnic Favoritism Within Constituencies

We next evaluate whether ethnic favoritism within constituencies increases with segregation (**H2**). We start with an intuitive yet powerful way of testing this hypothesis, using a set of difference-in-differences. We examine the 3502 localities in 120 constituencies that were *not* ethnically matched with their MP prior to 1998 (based on parliamentary elections in 1994). In the 1999 and 2004 elections, 55 of those 120 constituencies experienced a change in the ethnicity of their MP, resulting in 1599 localities becoming ethnically matched (*treated*) and 1903 localities remaining unmatched (*control*). We divide localities into low, medium, and highly segregated constituencies using the same terciles of the spatial dissimilarity index as above. For each level of segregation, we calculate the proportion of localities that have at least one borehole, by time period (1998 versus

[16]Throughout, we generate expected values and confidence intervals based on 10,000 simulations that approximate the sampling distribution of the parameters in the model (King, Tomz, and Wittenberg 2000; Gelman and Hill 2006, Ch. 7).

[17]Segregation might be particularly associated with public goods provision in electorally competitive constituencies. Because there are few competitive constituencies and we cannot attribute borehole to specific legislative terms, we are limited in our ability to evaluate this expectation. However, Table SI.12 and Figure SI.6 in the SI provide suggestive evidence that the predicted patterns are most pronounced in electorally competitive areas, although the differences are not statistically significant.

Table 1: Segregation and Borehole Investments across Electoral Constituencies

	Dependent variable:	
	Number of New Boreholes	
	(1)	(2)
Segregation (continuous)	1.31	
	(0.72)	
Dummy for Medium Segregation		0.43
		(0.14)
Dummy for High Segregation		0.36
		(0.16)
Ethnic Diversity (ELF)	−0.91	−0.86
	(0.38)	(0.37)
Population (10,000s)	0.07	0.06
	(0.02)	(0.02)
Population Density	−0.62	−0.71
	(0.25)	(0.23)
No. Boreholes, 1998	0.04	0.04
	(0.01)	(0.01)
Intercept†	2.71	3.02
	(0.54)	(0.42)
Admin. District FE	*Yes*	*Yes*
Observations	183	183

†Median administrative district intercept (Chikwawa) displayed

2008) and treatment status (control versus treated). Thus, we are simply comparing four means at each level of segregation. This difference-in-differences approach allows us to hold constant any time-invariant locality characteristics that influence public goods provision, including local ethnic diversity, collective-action capacity, and locality demand for public goods.

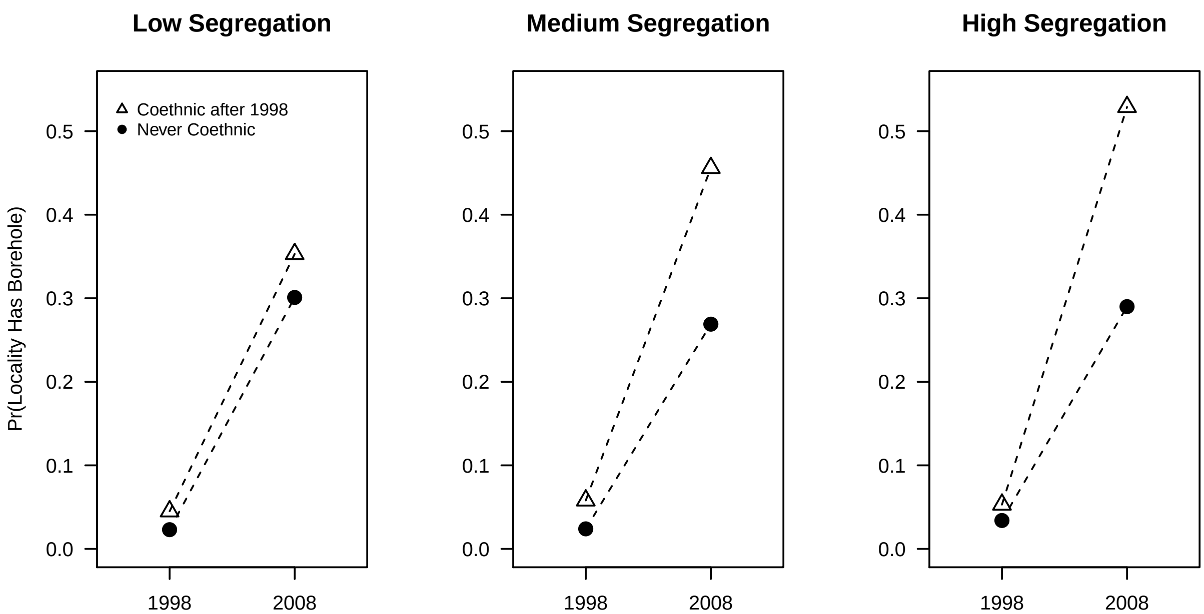
The results in Figure 4 provide evidence that segregation shapes whether MPs engage in ethnic favoritism in the allocation of boreholes. At low levels of segregation, there is no evidence of ethnic favoritism: 4.5 percent of treated and 2.3 percent of control localities had a borehole prior to 1998, which increased to 35.3 percent (treated) and 30.1 percent (control) by 2008. The difference-in-differences of 2.9 percentage points is not statistically significant ($p = 0.24$).[18] In contrast, we find evidence of ethnic favoritism in moderately and highly segregated constituencies. While treated and control localities had similar levels of borehole provision prior to 1998, localities that became ethnically matched with their MP after 1998 had a much higher chance of receiving a new borehole than localities that remained unmatched. The difference-in-differences is 15.2 percentage points ($p < 0.001$) in constituencies with medium levels of segregation, and 22.1 percentage points ($p < 0.001$) in constituencies with high levels of segregation. We can reject the null hypothesis that the difference-in-differences for low-segregation constituencies is the same as the difference-in-differences for medium- and high-segregation constituencies ($p < 0.001$). In short, we find evidence of ethnic favoritism, but only in moderately and highly segregated constituencies.[19]

We next evaluate whether this positive relationship between segregation and ethnic favoritism holds in the full sample of localities. We estimate the probability that a locality i in constituency c received at least one new borehole between 1998 and 2008 using the following

[18]Two-tailed test with standard errors clustered on localities to account for the panel structure of the data.

[19]We present a range of robustness tests of these results in the SI, including alternative segregation cutpoints (Figure SI.4) and a parametric test of how ethnic favoritism varies as a continuous function of segregation (Figure SI.3).

Figure 4: Ethnic Favoritism Is More likely in Segregated Constituencies



Note: Analysis includes 3502 localities nested in 120 constituencies. All of these localities were *not* coethnic with their MP in 1998. 1599 became coethnic with their MP in either the 1999 or the 2004 parliamentary elections; these are denoted with a triangle. The 1903 localities denoted with a circle were never coethnic with their MP in the study period.

model:

$$\Pr(Y_i = 1) = \Lambda\left\{\alpha_{c[i]} + \beta Match_i + \delta(Match_i \cdot Seg_{c[i]}) + X_i'\gamma\right\} \tag{2}$$

where $\Lambda\{\cdot\}$ is the cumulative distribution function of the logistic distribution. Since individual MPs only influence allocations within their constituency, we include constituency fixed effects, represented by $\alpha_{c[i]}$, which ensures that our estimates are driven by within-constituency variation.[20] *Binary Match* is a dummy variable equal to 1 if a locality's largest group was ethnically matched

[20] A conditional maximum likelihood estimator due to Chamberlain (1980) can be used to consistently estimate the structural parameters in fixed effects logit models. The Chamberlain approach does not estimate the incidental parameters (the constituency fixed effects), meaning that predicted

with the MP in at least one of the two legislative terms between 1998 and 2008, while *Continuous Match* is the share of a locality's population that is coethnic with the MP. To test whether ethnic favoritism varies by segregation, we interact *Match* with the constituency's level of segregation (*Seg*).[21] This measure of segregation is continuous, which forces us to assume a linear partial relationship (on the logit scale) between borehole investments and the interaction between ethnic favoritism and segregation. To relax this assumption, we also consider a model in which we interact *Match* with two dummy variables indicating medium and high degrees of segregation. To adjust for confounding, we include measures of the locality's population size (in 1000s), population density (in 1000s per square kilometer), and the number of boreholes the locality had in 1998 in vector X_i. Controlling for population density and boreholes in 1998 help us to account for potential differences in demand across localities. We cluster the standard errors at the constituency level.

The results are reported in Table 2. Model 1, which uses the continuous measure of segregation and binary match, is consistent with H2. At the lowest level of segregation observed in the data (dissimilarity score of 0.04), we expect ethnically matched localities to have a 44 percent chance of receiving a new borehole, holding covariates at their mean or mode value. The probability that ethnically matched localities received a new borehole was 10 percentage points higher in constituencies with median levels of segregation, and an additional 4 points higher in constituen-

probabilities and partial effects cannot be computed. The bias in the unconditional estimator is not severe when the number of observations within each cluster reaches about 10 (Katz 2001; Greene 2012, Ch. 17), which is the case for all but one of our constituencies. The conditional logit estimates presented in Table SI.6 in the SI are therefore almost identical to the estimates we present in Table 2. Table SI.6 also presents linear probability models estimated using OLS.

[21] Note that while we cannot identify the baseline effect of segregation because of perfect collinearity with the constituency fixed effects, we *can* estimate how segregation conditions the effect of *Match*. See Wooldridge (2002, Ch. 11) on interacting variables that are perfectly collinear with a set of fixed effects.

Table 2: Segregation and Within-Constituency Targeting of Boreholes

	Dependent variable:			
	I(EA Received Borehole)			
	Binary Match		Continuous Match	
	(1)	(2)	(3)	(4)
Ethnic Match	−0.04	0.001	0.14	0.32
	(0.38)	(0.20)	(0.98)	(0.39)
Ethnic Match x Segregation	0.72		1.27	
	(0.87)		(2.11)	
Ethnic Match x Med. Segregation		0.64		0.73
		(0.26)		(0.55)
Ethnic Match x High Segregation		0.09		0.31
		(0.29)		(0.57)
Population (1000s)	0.72	0.72	0.70	0.70
	(0.08)	(0.08)	(0.08)	(0.08)
Population Density	−0.30	−0.30	−0.29	−0.29
	(0.10)	(0.10)	(0.09)	(0.09)
No. Boreholes, 1998	−0.67	−0.67	−0.61	−0.61
	(0.09)	(0.09)	(0.11)	(0.11)
Intercept†	−1.05	−0.83	−1.08	−1.01
	(0.15)	(0.23)	(0.13)	(0.21)
Constituency FE	*Yes*	*Yes*	*Yes*	*Yes*
No. Localities (EAs)	9482	9482	9752	9752
No. Constituencies	148	148	149	149

†Median constituency intercept (Machinga North East) displayed
Note: Only constituencies with at least one new borehole are included.

cies with the highest levels (90^{th} percentile) of segregation. However, these estimates should be interpreted with caution due to the relatively large standard error on the segregation coefficient. Model 2 indicates that the positive association in Model 1 is driven by localities in constituencies with medium levels of segregation. According to Model 2, whereas ethnically matched localities in integrated constituencies had a 39 percent chance of receiving a new borehole, ethnically matched localities in medium-segregation constituencies had a 56 percent chance of receiving a new borehole. The same number for highly segregated constituencies is 44 percent. The difference between integrated and medium-segregation constituencies is statistically significant at conventional levels of confidence.[22] Models 3 and 4 present results from similar analyses using *Continuous Match*.[23] The results are qualitatively similar, although they are estimated less precisely: the interaction between continuous ethnic match and the continuous (Model 3) and categorical (Model 4) measures

[22]In 2006, a CDF was introduced that provided MPs with resources to sponsor development projects in their constituencies. We do not leverage the establishment of the CDF in the main analyses because we do not know whether boreholes were built before or after 2005. It is still possible, however, that ethnic match with the MP is more important after the introduction of the CDF. Table SI.9 in the SI show that our results are robust to measuring ethnic match separately by legislative term, but also that ethnic favoritism increased in moderate and high segregation constituencies following the CDF's establishment.

[23]We believe the binary measure is most appropriate because the continuous measure imposes linearity, assuming for example that the effect of a change from 10 to 20% coethnics is the same as a change from 45 to 55% coethnics. We also anticipate that MPs do not have the ability to make such fine grained distinctions among localities within their constituencies, but will have a general sense of where their own group dominates (which is captured in the binary indicator).

of segregation are positive, and ethnic favoritism in borehole provision is highest in moderately segregated constituencies.[24]

While we prefer the DiD analysis, which controls for time-invariant differences in localities' probability of receiving a new borehole and for common time shocks across localities (at a given level of segregation), these cross-sectional results increase our confidence that segregation conditions ethnic favoritism across a large number of localities. Coupled with our constituency-level results, there is substantial empirical support for our theoretical framework: segregation shapes both investment levels and ethnic favoritism with respect to the geographic allocation of local public goods.

Generalizing to Other Public Goods

To ensure that our conclusions are not limited to a single good (Kramon and Posner 2013), we also evaluate the impact of segregation on the provision of two other public goods — health clinics and schools.[25] The results in Table 3, which also inlcudes our borehole results for comparison, shows that segregation predicts the provision of all three goods across and within constituencies.[26]

These analyses also allow us to evaluate an extension of the theory. Schools and clinics differ from boreholes in terms of their greater reach. Thus, more coethnics likely benefit from the provision of these goods under high segregation than under moderate segregation (unlike with boreholes). We therefore expect segregation's impact on these goods to keep rising beyond moderate levels. Table 3 provides evidence for this expectation. Segregation's impact on borehole provision levels out at medium levels of segregation: the predicted number of new boreholes is

[24]Despite data limitations, we report some evidence in Table SI.13 that the effect of ethnic match, conditional on segregation, is more pronounced in electorally competitive constituencies.

[25]Data on these goods were also collected during the 1998 and 2008 censuses.

[26]Full results presented in Tables SI.10 and SI.11.

about 30 at both medium and high levels. By contrast, there is a marked (though not always statistically significant) difference in the provision of clinics and schools across medium and high levels of segregation. The predicted number of new schools is 3 (0.34 of a standard deviation in the number of new schools) at low levels of segregation, 5 (0.6) at medium levels, and 7 (0.8) at high levels. The same numbers for clinics are 1.2 (0.64), 1.6 (0.85), and 2 (1.1). Similarly, the effect of segregation on ethnic targeting within constituencies continues to increase beyond moderate levels for schools and clinics, but not for boreholes (Models 4-6). Taken together, these analyses suggest that the reach of public goods impacts the relationship between segregation and their provision.

Table 3: Segregation's Effect on Boreholes, Clinics, and Schools

	Dependent variable:					
	Number of New Goods			I(EA Received New Good)		
	Boreholes	Clinics	Schools	Boreholes	Clinics	Schools
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy for Medium Segregation	0.43	0.30	0.56			
	(0.14)	(0.23)	(0.18)			
Dummy for High Segregation	0.36	0.49	0.83			
	(0.16)	(0.27)	(0.22)			
Ethnic Match				0.001	−0.87	−0.08
				(0.20)	(0.28)	(0.21)
Ethnic Match x Med. Segregation				0.64	1.23	0.06
				(0.26)	(0.47)	(0.32)
Ethnic Match x High Segregation				0.09	2.77	0.46
				(0.29)	(0.84)	(0.33)
Other Controls & Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
No. Constituencies	183	183	183	148	86	138
No. Localities (EAs)				9482	5735	8954

Notes: Controls are the same as in Tables 1 and 2, except *No. Boreholes, 1998* is replaced by clinic and school equivalents in models 2-3 and 5-6. Only constituencies with at least one new good are included in the within-constituency models (4-6).

Alternative Explanations

While the empirical patterns reported above are consistent with our theory, this section discusses and empirically assesses a number of alternative explanations.

Local Ethnic Homogeneity and Collective Action. One alternative centers on the expectation that homogenous localities are better able to collectively mobilize to locally produce public goods (Miguel and Gugerty 2005; Habyarimana et al. 2009). This poses a potential challenge to our interpretation of the constituency-level results because segregated constituencies will tend to have more homogenous localities than integrated ones. If public goods are locally produced at a higher rate in homogenous localities, then segregated constituencies would mechanically have more public goods. This explanation is, however, inconsistent with our locality-level results (Figure 4), which show that only ethnically matched localities receive more public goods in segregated constituencies. If local ethnic homogeneity alone were driving our results, we would not expect the effect of segregation to be conditional on ethnic match with an MP.

MP Quality. Another alternative explanation is that some unobserved characteristics of constituencies are correlated with the degree of segregation and the quality of the MP, producing a spurious relationship between segregation and public goods provision. To address this concern, we carry out a placebo test that examines whether segregation affects the provision of *private goods* in the form of agricultural subsidies. Like local public goods, agricultural subsidies are highly valued by Malawians (Harrigan 2008). But unlike local public goods, they are targeted to specific households, meaning that segregation should be less consequential for their provision. If segregation were positively associated with the provision of these goods, this would suggest that MP quality might be driving our constituency-level results. Examining survey responses from 11,000 households on the largest and most politically salient form of private transfers from the Malawian government to citizens — coupons that subsidize agricultural inputs through the Targeted Input Program — we find that the provision of these goods is not affected by segregation, or may even be decreasing with segregation (Table SI.14). Thus, MP quality is unlikely to account for the relationship between segregation and public goods provision.

Residential Sorting. If Malawians move in response to the provision of public goods, then our ability to detect the effect of ethnic demography on their provision could be threatened. However, we anticipate that such residential sorting would lead to more diverse populations, and thus more integration, near public goods, as migrants move towards better served areas — the opposite pattern of what we observe. Even if residential sorting could account for the positive association between ethnic segregation and public goods, it would also need to account for a greater ability to elect a coethnic MP in order for this to threaten our conclusions. Furthermore, we anticipate that rural-rural migration in Malawi is relatively constrained due to the scarcity of land and customary rules governing land tenure (Chirwa 2008; Government of Malawi 2001; Kishindo 2004). Census data shows in fact that only 10% of rural Malawians reside outside their district of birth.[27] And what rural-rural migration does exist is unlikely to shift the ethnic landscape because both marriage and accessing communally held land typically occur within ethnic communities.[28] Rural-rural migration across ethnic communities is typically limited to laborers on large tobacco or tea estates (Potts 2006), areas which are likely to have more, not less, public goods provision. An important exception to this general pattern was a large scale land resettlement program that moved over 15,000 families from overpopulated areas of Malawi to fallow estates in Mangochi and Machinga between 2005 and 2011 (Chinsinga 2011). To ensure that this is not influencing our results, we show in Columns 5 and 6 of Tables SI.4 and SI.5 of the SI that our main results are robust to

[27] This figure is based on individual-level information about district of birth and district of residence in a random 10% sample ($n = 1,282,335$) of the 2008 census data, accessed from the IPUMS dataset (Minnesota Population Center 2014).

[28] There are customary and cultural barriers to rural migrants accessing land outside their ethnic community (Potts 2006), and most marriages within Malawi are formed within 5 miles of one's home village (reported in Englund 2002).

removing Machinga and Mangochi constituencies. Taken together, these patterns of migration and robustness tests suggest that residential sorting is highly unlikely to drive our results.

Differences in Demand. Another possibility is that differential demand for boreholes is driving the results. We address this in a number of ways. First, because access to clean water is a basic need, the strongest predictor of demand will be current access. By controlling for the number of boreholes present in 1998, we control for reduced demand resulting from already having a borehole. Second, we emphasize that an advantage of the DiD analysis (Figure 4) is that it effectively controls for fixed unobserved differences between localities, including their demand for boreholes and other local public goods. For differential demand to account for the DiD results, newly matched localities would have to experience greater increases in demand for clean water than localities who remained unmatched *and* this differential increase in demand would have to occur only in segregated constituencies, which seems unlikely. Third, we deal with variation in demand that is driven by access to other water sources, in particular the fact that many urban dwellers have access to piped water. In our main analyses, we control for population density (a strong proxy for urban areas), which is likely to capture reduced demand. In addition, we show that our results are robust to removing urban constituencies from the analysis (Tables SI.4 and SI.5 of the SI).

Plurality Group Favoritism. Finally, we have interpreted our results as evidence of in-group favoritism. It is possible, however, that MPs are instead targeting benefits to the largest ethnic group in a constituency, whether it is their own group or not, in order to maximize their electoral coalition. With few constituencies in which the MP is not a member of the ethnic plurality, we cannot distinguish plurality group favoritism from coethnic favoritism. We note, however, that the interpretation we have offered is plausible in light of the existing evidence that politicians in much of Africa have incentives to favor their own ethnic group. Furthermore, this alternative interpretation does not undermine our general argument: regardless of the group that the political elite is seeking to favor, our logic suggests that the segregation of that group shapes how it is favored.

Conclusion

This paper advances a theory about how ethnic segregation shapes elite strategies for engaging in ethnic favoritism. We show that MPs in Malawi allocate more public goods to their constituencies when ethnic groups are spatially segregated, and that ethnic favoritism in public goods provision is more common in segregated contexts. These patterns are consistent with our claim that ethnic segregation conditions how elites invest in and allocate public goods.

Our theory and results make several contributions to the study of ethnic politics in Africa and distributive politics more broadly. First, they underscore the importance of ethnic segregation when studying distributive politics in diverse contexts. While scholars have long noted the geographic clustering of ethnic groups in Africa (e.g., Bates 1983; Kimenyi 2006), there has been little systematic evaluation of how such segregation influences material outcomes. This is in stark contrast to the well-documented effects of segregation on intergroup attitudes, trust, governance, and political participation (e.g., Alesina and Zhuravskaya 2011; Enos 2014; Kasara 2013; Key 1949; Oliver and Wong 2003; Robinson 2015).

Our framework also helps make sense of outstanding puzzles in the empirical literature on ethnic politics in Africa. For example, while ethnic favoritism is pervasive in some contexts, it is not universal (Franck and Rainer 2012). Nor is there ethnic favoritism in the allocation of all distributive goods within a given context (Kramon and Posner 2013). Our theory contributes by specifying the conditions under which ethnic favoritism should manifest in local public goods provision. Furthermore, because Africa is marked by quite high levels of ethnic segregation relative to other regions, our theory may help account for the high degree of ethnic favoritism in public goods provision there. Our theory also has direct implications for the question of why local ethnic diversity is often associated with low public goods provision. While past explanations focus on local collective action (Alesina, Baqir, and Easterly 1999; Habyarimana et al. 2009; Miguel and Gugerty 2005), our framework suggests that political leaders under-invest in public goods in highly diverse local areas because such goods are too difficult to target to their coethnic supporters. Thus, distributive politics may help to account for the under-provision of public goods in ethnically diverse areas.

Our study also contributes to recent work on ethnic geography and vote choice. While we do not observe vote choice in Malawi, our theory implicitly generates expectations about the relationship between ethnic segregation and ethnic-based voting. Past research has found that the geographic concentration of ethnic groups is positively associated with ethnic bloc voting in Africa, but attributed this to proximity and common preferences (Ishiyama 2012). Our results suggest that geographically segregated groups will tend to vote ethnically because they anticipate that local public goods will be targeted to their area. Consistent with this expectation, Nathan (2015) finds that variation in ethnic segregation across urban neighborhoods in Ghana predicts ethnic voting, which he attributes to the (unobserved) expectation that politicians provide different types of goods to localities with different ethnic geographies. In rural Ghana, Ichino and Nathan (2013) find that citizens who make up a local ethnic minority are willing to vote for a non-coethnic presidential candidate in expectation of receiving local public goods. Our study is consistent with such voter expectations, but also implies that local ethnic minorities should be most likely to vote across ethnic lines in contexts of high ethnic segregation.

Finally, while we test the theory in Malawi, we expect the argument to generalize to other contexts for two reasons. First, Malawi is similar to many other countries in that political elites have incentives to favor some groups over others. Research on distributive politics shows this to be the case in a range of socio-economic and institutional contexts: in Australia, a wealthy democracy with single-member districts (Denemark 2000); in Sweden, a wealthy democracy with proportional representation (Dahlberg and Johansson 2002); in India, a developing democracy with single-member districts (Min 2015); in Benin, a developing democracy with proportional representation (Kramon and Posner 2013); and in Egypt, an electoral authoritarian regime (Blaydes 2010). Second, because our theory emphasizes the importance of segregation in shaping the type of goods used to favor one group over others, the theory can be applied to the study of favoritism in contexts where elites have discretion over different types of goods (private and public). In urban Ghana, for example, Nathan (2015) finds that voters expect elites to distribute different types of goods to neighborhoods with different ethnic demographies, which is consistent with our framework. Research from Latin America documents that governments often invest in a different mix of public and private goods in different local political contexts (Albertus 2012; Magaloni, Diaz-Cayeros, and Estévez 2007), patterns that our logic may help to explain. Thus, while more research is required, we anticipate that segregation may shape distributive politics in contexts with different institutional configurations, degrees of urbanization, and levels of economic development. In short, our central finding — that ethnic segregation conditions the strategies that incumbents use to favor their coethnics — has implications for the study of distributive politics beyond Africa. Wherever political elites have incentives to favor certain groups of voters over others, the spatial distribution of these groups is likely to shape the distributive strategies they adopt.

References

Albertus, Michael. 2012. "Vote Buying With Multiple Distributive Goods." *Comparative Political Studies* 46(9): 1082–1111.

Alesina, Alberto, and Ekaterina Zhuravskaya. 2011. "Segregation and the Quality of Government in a Cross Section of Countries." *American Economic Review* 101(5): 1872–1911.

Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics* 114(4): 1243–1284.

Bates, Robert. 1983. "Modernization, Ethnic Competition, and the Rationality of Politics in Contemporary Africa." In *State versus Ethnic Claims: African Policy Dilemmas*, ed. Donald Rothchild, and Victor A. Olorunsola. Boulder, CO: Westview Press pp. 152–71.

Baumann, Erich, and Kerstin Danert. 2008. "Operation and Maintanance of Rural Water Supplies in Malawi." SKAT Report. Available at: `http://www.rural-water-supply.net/_ressources/documents/default/208.pdf`.

Blaydes, Lisa. 2010. *Elections and Distributive Politics in Mubarak's Egypt*. New York: Cambridge University Press.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padro i Miquel. 2015. "The Value of Democracy: Evidence from Road Building in Kenya." *American Economic Review* 105(6): 1817–51.

Cammack, Diana, Fred Golooba-Mutebi, Fidelis Kanyongolo, and Tam ONeil. 2007. "Neopatrimonial Politics, Decentralisation and Local Government: Uganda and Malawi in 2006." In *Good Governance, Aid Modalities and Poverty Reduction*. Research project (RP-05-GG): Advisory Board for Irish Aid.

Carlson, Elizabeth. 2015. "Ethnic Voting and Accountability in Africa: A Choice Experiment in Uganda." *World Politics* (February): 1–33.

Chamberlain, Gary. 1980. "Analysis of Covariance with Qualitative Data." *The Review of Economic Studies* 47: 225–238.

Chasukwa, Michael, Asiyati L Chiweza, and Mercy Chikapa-Jamali. 2014. "Public Participation in Local Councils in Malawi in the Absence of Local Elected Representatives-Political Eliticism or Pluralism?" *Journal of Asian and African Studies* 49(6): 705–720.

Chinsinga, Blessings. 2005. "District Assemblies in a Fix: The Perils of Self-Seeking Tendencies in Decentralisation Policy Reforms in Malawi." *Development Southern Africa* 22(4): 529–548.

Chinsinga, Blessings. 2007. *Democracy, Decentralisation, and Poverty Reduction in Malawi*. Cologne, Germany: Rüdiger Köppe Verlag.

Chinsinga, Blessings. 2009. "The Interface Between Local Level Politics, Constitutionalism and State Formation in Malawi Through the Lens of Constituency Development Fund (CDF)." Available at: `http://www.eldis.org/go/home&id=58474&type=Document#.VPhvnEI1Bds`.

Chinsinga, Blessings. 2011. "The politics of land reforms in Malawi: The case of the Community Based Rural Land Development Programme (CBRLDP)." *Journal of International Development* 23(3): 380–393.

Chirwa, Ephraim. 2008. "Land Tenure, Farm Investments, and Food Production in Malawi." IPPG Paper No. 18. Available at: `http://www.eldis.org/go/home&id=58514&type=Document#.VZcXRs7seVg`.

Conroy-Krutz, Jeffrey. 2012. "Information and Ethnic Politics in Africa." *British Journal of Political Science* 43(345-73).

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

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

DeGabriele, Joseph. 2002. "Improving Community Based Management of Boreholes: A Case Study from Malawi." Land Tenure Center, University of Wisconsin-Madison. Available at: `http://agris.fao.org/agris-search/search.do?recordID=GB2013202862`.

Denemark, David. 2000. "Partisan Pork Barrel in Parliamentary Systems: Australian Constituency Level Grants." *Journal of Politics* 62(3): 896–915.

Dixit, Avinash, and John Londregan. 1996. "The Determinants of Success of Special Interests in Redistributive Politics." *Journal of Politics* 58(4): 1132–1155.

Ekeh, Peter P. 1975. "Colonialism and the two publics in Africa: A theoretical statement." *Comparative Studies in Society and History* 17(01): 91–112.

Englund, Harry. 2002. "The Village in the City, the City in the Village: Migrants in Lilongwe." *Journal of Southern African Studies* 28(1): 137–154.

Enos, Ryan D. 2014. "Causal Effect of Intergroup Contact on Exclusionary Attitudes." *Proceedings of the National Academy of Sciences* 111(10): 3699–3704.

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.

Gelman, Andrew, and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge, UK: Cambridge University Press.

Gilman, Lisa. 2009. *The Dance of Politics: Gender, Perfromance, and Democratization in Malawi*. Philadelphia, PA: Temple University Pres.

Golden, Miriam, and Brian Min. 2013. "Distributive politics around the world." *Annual Review of Political Science* 16: 73–99.

Government of Malawi. 1998. "Population and Housing Census." National Statistical Office, Zomba, Malawi.

Government of Malawi. 2001. "Malawi National Land Policy." Ministry of Lands and Housing, Lilongwe, Malawi.

Government of Malawi. 2008. "Population and Housing Census." National Statistical Office, Zomba, Malawi.

Greene, William. 2005. "Functional Form and Heterogeneity in Models for Count Data." *Foundations and Trends in Econometrics* 1(2): 113–218.

Greene, William H. 2012. *Econometric Analysis*. La Jolla, CA: Granite Hill Publishers.

Habyarimana, James P., Macartan Humphreys, Daniel N. Posner, and Jeremy Weinstein. 2009. *Coethnicity: Diversity and the Dilemmas of Collective Action*. New York: Russell Sage Foundation Publications.

Harrigan, Jane. 2008. "Food Insecurity, Poverty and the Malawian Starter Pack: Fresh Start or False Start?" *Food Policy* 33(3): 237–249.

Ichino, Nahomi, and Noah Nathan. 2013. "Local Ethnic Geography and Instrumental Ethnic Voting: Voter Behavior Across Rural Ghana." *American Political Science Review* 107(2): 344–61.

Ishiyama, John. 2012. "Explaining Ethnic Bloc Voting in Africa." *Democratization* 19(4): 761–788.

Joseph, Richard. 1987. *Democracy and Prebendalism in Nigeria*. New York, NY: Cambridge University Press.

Kasara, Kimuli. 2007. "Tax Me If You Can: Ethnic Geography, Democracy, and the Taxation of Agriculture in Africa." *American Political Science Review* 101(01): 159–172.

Kasara, Kimuli. 2013. "Separate and Suspicious: Local Social and Political Context and Ethnic Tolerance in Kenya." *Journal of Politics* 75(04): 921–936.

Katz, Ethan. 2001. "Bias in Conditional and Unconditional Fixed Effects Logit Estimation." *Political Analysis* 9(4): 379–384.

Key, V.O. 1949. *Southern Politics in State and Nation*. New York: Knopf.

Kimenyi, Mwangi S. 2006. "Ethnicity, Governance and the Provision of Public Goods." *Journal of African Economies* 15: 62–99.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2): 347–361.

Kishindo, Paul. 2004. "Customary Land Tenure and the New Land Policy in Malawi." *Journal of Contemporary African Studies* 22(2): 213–225.

Kramon, Eric. 2013. *Vote Buying and Accountability in Democratic Africa*. PhD Dissertation, University of California, Los Angeles.

Kramon, Eric, and Daniel N. Posner. 2013. "Who Benefits from Distributive Politics? How the Outcomes One Studies Affect the Answers One Gets." *Perspectives on Politics* 11(2).

Lindberg, Staffan I. 2003. "It's Our Time to "Chop": Do Elections in Africa Feed Neo-Patrimonialism rather than Counter-Act It?" *Democratization* 10(2): 121–40.

Lindberg, Staffan I. 2010. "What accountability pressures do MPs in Africa face and how do they respond? Evidence from Ghana." *Journal of Modern African Studies* 48(01): 117–142.

Magaloni, Beatriz, Alberto Diaz-Cayeros, and Federico Estévez. 2007. "Clientelism and portfolio diversification: A model of electoral investment with applications to Mexico." In *Patrons, Clients, and Policies: Patterns of Democratic Accountability and Political Competition*, ed. Herbert Kitschelt, and Steven I. Wilkinson. New York: Cambridge University Press pp. 182–205.

Matuszeski, Janina, and Frank Schneider. 2006. "Patterns of Ethnic Group Segregation and Civil Conflict." Working Paper, Harvard University. Available at: http://www.cgdev.org/doc/events/02.09.07/Matuszeski-JMP.pdf.

Miguel, Edward, and Mary Kay Gugerty. 2005. "Ethnic diversity, Social Sanctions, and Public Goods in Kenya." *Journal of Public Economics* 89(11-12): 2325–2368.

Min, Brian. 2015. *Power and the Vote: Elections and Electricity in the Developing World*. New York: Cambridge University Press.

Minnesota Population Center. 2014. *Integrated Public Use Microdata Series, International: Version 6.3*. Minneapolis: University of Minnesota.

Nathan, Noah. 2015. "Local Ethnic Geography, Expectations of Favoritism, and Voting in Urban Ghana." Working Paper, Harvard University. Available at: http://scholar.harvard.edu/noahnathan/publications/local-ethnic-geography-expectations-favoritism-and-voting-urban-ghana.

Nichter, Simeon. 2008. "Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot." *American Political Science Review* 102(01): 19–31.

Oliver, Eric J, and Janelle Wong. 2003. "Intergroup Prejudice in Multiethnic Settings." *American Journal of Political Science* 47(4): 567–582.

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

Potts, Deborah. 2006. "Rural Mobility as a Response to Land Shortages: The Case of Malawi." *Population, Space and Place* 12: 291–311.

Reardon, Sean F, and David O'Sullivan. 2004. "Measures of Spatial Segregation." *Sociological Methodology* 34(1): 121–162.

Robinson, Amanda Lea. 2015. "Ethnic Diversity, Segregation, and Ethnocentric Trust in Africa." Presented at the International Studies Association Annual Meeting, New Orleans, LA. Available at: `https://u.osu.edu/robinson.1012/files/2015/03/Segregation_CETP-2cw4xlm.pdf`.

Stokes, Susan, Thad Dunning, Marcelo Nazareno, and Valeria Brusco. 2013. *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics*. Cambridge, UK: Cambridge University Press.

Tajfel, Henri, and John C. Turner. 1985. "The Social Identity Theory of Intergroup Behavior." In *Psychology of Intergroup Relations*, ed. Stephen Worchel, and William G. Austin. Chicago, Ill.: Nelson Hall.

The Malawi Voice. 2014. "Dr Jean Kalilani Rating High in Dowa." November 14.

The Nation. 2012. "MP commissions K1.6m borehole in Zomba." July 11.

van de Walle, Nicolas. 2003. "Presidentialism and clientelism in Africa's emerging party systems." *Journal of Modern African Studies* 41(2): 297–321.

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

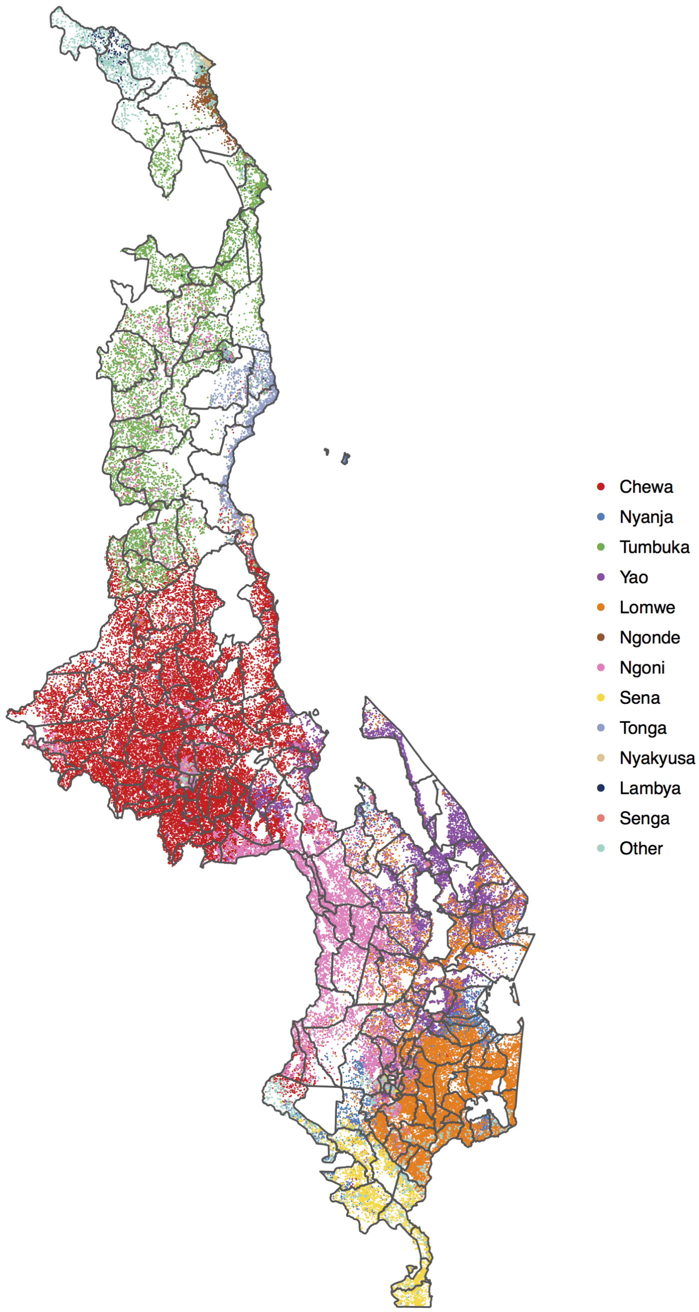
Wooldridge, Jeffrey M. 1997. "Quasi-Likelihood Methods for Count Data." In *Handbook of Applied Econometrics*, ed. M. Hashem Pesaran, and Peter Schmidt. Vol. 2 Oxford, UK: Blackwell pp. 352–406.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.

Supporting Information:
Segregation, Ethnic Favoritism, and the Strategic Targeting of Local Public Goods

Descriptive Statistics

Figure SI.1: The Spatial Distribution of Malawi's Major Ethnic Groups



Note: Each dot in the figure represents 100 individuals, and has been color coded according to ethnicity. The gray borders delineate Malawi's 193 electoral constituencies.

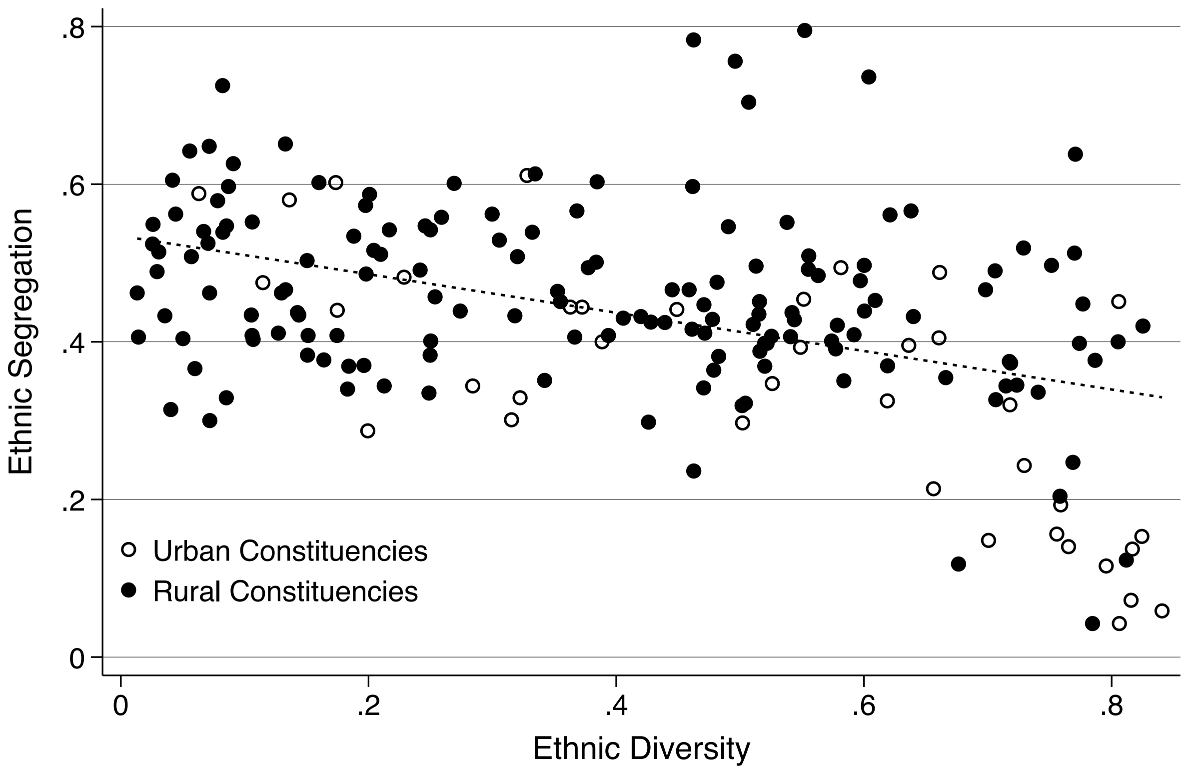
Table SI.1: Summary Statistics Across Localities

	Mean	SD	Min	Max	N
Demographics					
Population (in 1,000s)	1.04	0.55	0.00	8.29	12380
Population Density (in 1,000s/sqkm)	1.12	4.14	0.00	90.95	12380
Ethnic Group Majority Present	0.80	0.40	0.00	1.00	12380
Proportion of Largest Ethnic Group	0.74	0.24	0.00	1.00	12380
Ethnic Diversity (ELF)	0.32	0.26	0.00	0.87	12380
Representation					
MP Ethnic Match, 1999-2004	0.66	0.48	0.00	1.00	11953
MP Ethnic Match, 2004-2009	0.69	0.46	0.00	1.00	12035
MP Ethnic Match Ever, 1999-2009	0.76	0.43	0.00	1.00	11983
MP Ethnic Match Share, 1999-2004	0.59	0.37	0.00	1.00	11834
MP Ethnic Match Share, 2004-2009	0.60	0.37	0.00	1.00	12035
MP Ethnic Match Share Avg., 1999-2009	0.59	0.36	0.00	1.00	12292
Public Goods					
No. of Boreholes, 1998	0.05	0.32	0.00	6.00	12380
New Borehole Indicator, 1998-2008	0.33	0.47	0.00	1.00	12380
No. of New Boreholes, 1998-2008	0.61	1.10	0.00	13.00	12380
No. of Schools, 1998	0.33	0.59	0.00	6.00	12380
New School Indicator, 1998-2008	0.08	0.27	0.00	1.00	12380
No. of New Schools, 1998-2008	0.09	0.35	0.00	4.00	12380
No. of Clinics, 1998	0.05	0.24	0.00	4.00	12380
New Clinic Indicator, 1998-2008	0.02	0.13	0.00	1.00	12380
No. of New Clinics, 1998-2008	0.02	0.15	0.00	4.00	12380
Private Goods					
Proportion Receiving Fertilizer Subsidy, 2004	0.54	0.28	0.00	1.00	564

Table SI.2: Summary Statistics Across Electoral Constituencies

	Mean	**SD**	**Min**	**Max**	**N**
Demographics					
Population (in 10,000s)	6.69	3.15	0.42	22.99	193
Population Density (in 10,000s/sqkm)	0.41	1.00	0.01	9.10	193
Urban Constituency	0.09	0.28	0.00	1.00	193
Ethnic Diversity (ELF)	0.41	0.24	0.01	0.84	193
MP Ethnic Group Segregation	0.43	0.13	0.04	0.80	193
Public Goods					
No. of Boreholes, 1998	3.48	4.67	0.00	24.00	193
New Borehole Indicator, 1998-2008	0.82	0.38	0.00	1.00	193
No. of New Boreholes, 1998-2008	39.00	40.64	0.00	207.00	193
No. of Schools, 1998	20.95	14.22	0.00	64.00	193
New School Indicator, 1998-2008	0.69	0.46	0.00	1.00	193
No. of New Schools, 1998-2008	5.64	9.03	0.00	68.00	193
No. of Clinics, 1998	3.26	2.88	0.00	17.00	193
New Clinic Indicator, 1998-2008	0.44	0.50	0.00	1.00	193
No. of New Clinics, 1998-2008	1.16	1.93	0.00	15.00	193
Private Goods					
Proportion Receiving Fertilizer Subsidy, 2004	0.60	0.18	0.05	1.00	159

Figure SI.2: Relationship Between Diversity and Segregation Across Electoral Constituencies



Note: This figure shows the relationship between a constituency's degree of ethnic diversity, measured using the standard ethnolinguistic fractionalization index, and the degree of ethnic group segregation, measured using the average MP-specific ethnic group spatial dissimilarity index. The correlation coefficient is -0.43 across all 193 constituencies, but only -0.28 among rural constituencies. This negative relationship is driven by the fact that segregation is increasingly difficult as diversity increases. Despite the negative correlation, there is considerable variation in segregation at all levels of diversity.

Ethnicity Data for Members of Parliament

We compiled new data on the ethnic identity of each Malawian MP who served between 1994 and 2009. To assemble this dataset, we first gathered the names of all MPs from official election returns (Government of Malawi 1994; 1999; 2004). Two Malawian research assistants then coded the ethnic identity of each MP with the assistance of staff at the Malawi Electoral Commission, Administrative District Offices, and local elites within each constituency. The inter-rater reliability score across the two coders was 0.66. Where codings differ, we use the coding with the best documented sources.

Measure of Ethnic Group Segregation

Our fine-grained ethnicity data enable us to improve upon past efforts to examine the impact of segregation on ethnic favoritism. In particular, we improve upon the analysis in Franck and Rainer (2012), which finds no evidence that ethnic favoritism by African presidents is conditioned by country-level segregation. Franck and Rainer's measure of segregation comes from Matuszeski and Schneider (2006), who developed it from the spatial distribution of language groups provided in the Global Mapping International (GMI) dataset. GMI partitions the globe into mutually exclusive language-group polygons such that each area of the world has only one language group whose boundaries are defined and non-overlapping. Thus, the data cannot capture ethnic integration that occurs from members of more than one group living in the same local area. As our data show, however, there exists a great deal of local ethnic heterogeneity. Additionally, because Matuszeski and Schneider measure the segregation of language groups relative to a spatial grid within each country, levels of segregation on this measure are driven almost entirely by the number of ethnic group *borders* in a country: where there are more borders — because of more groups or because large groups reside in segregated pockets — the country is scored as less segregated. As a result, the measure is likely to generate misleading codings of segregation. Our disaggregated data thus allow for a more appropriate test of segregation's impact.

Using this fine-grained data, we rely on the spatial dissimilarity index to measure how geographically clustered different ethnic groups are relative to what an even geographic distribution of the ethnic groups would look like. This and its non-spatial counterpart are widely used and accepted measures of ethnic and racial segregation (e.g., Cutler, L., and Vigdor 2012). The non-spatial version of the dissimilarity index captures the deviation between locality and constituency ethnic group proportions. In the case of complete integration, all localities would have the same ethnic group proportions as the whole constituency. The spatial version is similar but also accounts for the ethnic make-up of neighboring localities. It measures the deviation between the ethnic composition of the constituency and individuals' "local environment," where the local environment can consist of (parts of) several neighboring localities. We implement this measure in `R`, using the function `spseg` in package `seg`.

In this section, we describe the theory behind the spatial dissimilarity measure in further detail.[1] We are interested in measuring the spatial distribution of two mutually exclusive groups: coethnics of the MP and non-coethnics of the MP. Index these two groups $g \in \{c, nc\}$ (c = coethnics; nc = non-coethnics). Further, let p index geographical locations within the MP's jurisdiction, which is denoted J, and let q index points located some distance from p. Parameters super-positioned with ˜ describe the local spatial environment of a point rather than the point itself. Table SI.3 describes each component of the spatial dissimilarity measure. In the table, "population density" means the population count per unit area (e.g., 10 m^2) at location p; $\phi_p = \int_{q \in J} \exp(-2||p-q||)dq$; and $||p-q||$ represents the euclidean distance between p and q. The spatial dissimilarity index $\tilde{D}$ is then:

$$\tilde{D} = \sum_g \int_{p \in J} \frac{\tau_p}{2NI} |\pi_p^g - \pi^g| dp$$

Note that this measure will approach 0, indicating minimum segregation, when the group proportions at the local environment (π_p^g) are similar to the overall ethnic composition of the jurisdiction (π^g). Further, $\exp(-2||p-q||)$ is one of many potential non-negative functions we could have chosen to define the proximity of p and q. In our case, the farther away p and q are, the less will q influence the local environment at p. This is the default option in `seg`, the `R` package we use to implement this measure.

[1]See Reardon and O'Sullivan (2004) for a detailed discussion of the nature and validity of this and other spatial segregation measures.

Table SI.3: Key Components of the Spatial Dissimilarity Index

Symbol	Concept	Equivalent expression
N	Total population in jurisdiction *J*	
τ_p	Population density at *p*	
τ_p^g	Population density of *g* at *p*	
$\tilde{\tau}_p$	Population density in local environment	$\frac{1}{\phi_p}\int_{q\in J}\tau_q \exp(-2\|\|p-q\|\|)dq$
$\tilde{\tau}_p^g$	Population density of *g* in local environment	$\frac{1}{\phi_p}\int_{q\in J}\tau_q^g \exp(-2\|\|p-q\|\|)dq$
π^g	Proportion of *g* of total population	
$\tilde{\pi}_p^g$	Proportion of *g* in local environment	$\frac{\tilde{\tau}_p^g}{\tilde{\tau}_p}$
I	Overall diversity of *J*	$2\pi^c\pi^{nc}$

Robustness Tests: Difference-In-Differences

The difference-in-difference (DiD) setup we use in the paper suggests that segregation shapes the degree to which politicians engage in ethnic favoritism (see Figure 4). The setup uses different cutoffs for segregation. Here, we carry out two sets of tests that demonstrate that the conclusions we draw are not dependent upon these cutoffs. The first test uses a regression model that does not depend on cutoffs. The second test carries out the same analysis as in the paper for 15 cutoffs (see Figure SI.4 on p. 12 of this appendix).

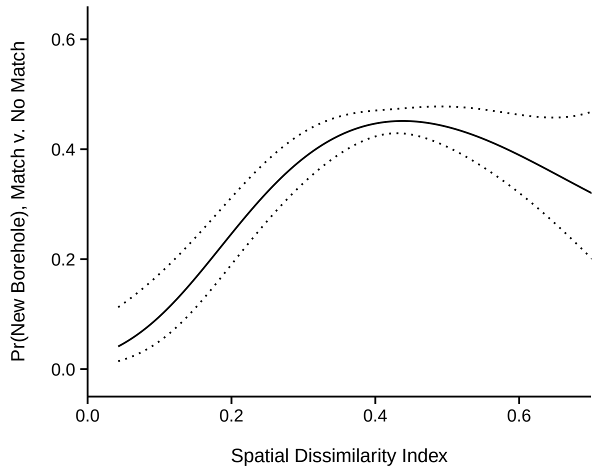
In the first test, the regression model predicting the receipt of a new borehole ($Y = 1$; 0 otherwise) looks as follows:

$$\Pr(Y_{icgt} = 1) = \Lambda\left\{\alpha + G'\gamma + P'\delta + D'\beta\right\}$$

$$\text{for} \quad G = \begin{bmatrix} g_g \\ g_g \cdot S_c \\ g_g \cdot S_c^2 \\ g_g \cdot S_c^3 \end{bmatrix} \quad P = \begin{bmatrix} p_t \\ p_t \cdot S_c \\ p_t \cdot S_c^2 \\ p_t \cdot S_c^3 \end{bmatrix} \quad D = \begin{bmatrix} d_{gt} \\ d_{gt} \cdot S_c \\ d_{gt} \cdot S_c^2 \\ d_{gt} \cdot S_c^3 \end{bmatrix}$$

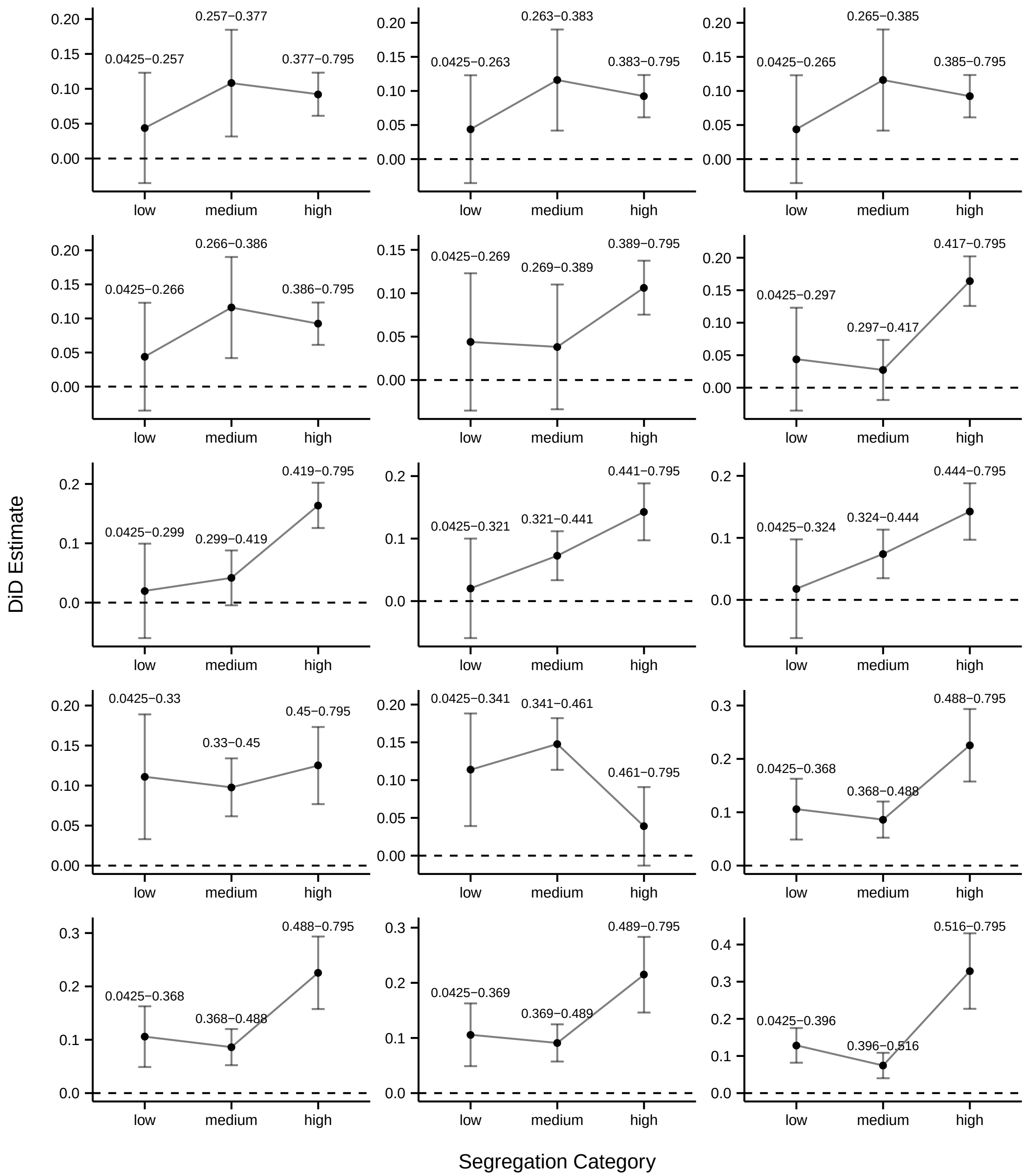
where i indexes locality, c constituency, g treatment group, t time period, and $\Lambda\{\cdot\}$ is the CDF of the logistic distribution. The variable g_g equals 1 for treated localities (those that became matched in the second period) and 0 otherwise, p_t equals 1 in the second period and 0 otherwise, d_{gt} equals 1 for treated localities in the second period and 0 otherwise, and S_c is a measure of segregation. This setup allows the DiD estimate to vary with segregation to a third-degree polynomial (captured by the vector β). We use a third-degree polynomial because we find significant evidence that fit is improved as compared to a second-degree polynomial or including no polynomial. Figure SI.3 shows the result, and aligns well with the results reported in the paper.

Figure SI.3: DiD Estimates as a Continuous Function of Segregation



Note: The y-axis is a measure of ethnic favoritism based on a difference-in-differences setup. For example, a 0.4 score on the y-axis indicates that the share of newly matched localities that received a new borehole in 1998-2008 was 40 percentage points higher than expected given the share of unmatched localities that received a new borehole. The figure provides parametric evidence that the DiD results we report in the paper are not sensitive to a particular definition of low, medium, and high segregation. For further details, see p. 10 of this appendix.

Figure SI.4: DiD Estimates For Different Segregation Cutoffs



Note: This figure shows 15 replications of Figure 4, but decomposes the four means for each level of segregation into one summary measure, the DiD (capturing ethnic favoritism). It then plots the DiD estimate for low, medium, and high segregation. Each subplot employs a different set of mutually exclusive cutoffs for segregation, randomly generated subject to the following constraints: the medium category lower cutoff has to fall in the interval [0.25, 0.4]; the high category lower cutoff is then set 0.12 points higher than the medium cutoff. This approach ensures that at least 10% of the data are included in each category.

Robustness Tests: Different Samples

In this section, we show that our constituency-level and within-constituency analyses are robust to changes in the sample that we analyze. First, we re-run our analyses using all constituencies. Recall that in the main analysis, we exclude all highly homogenous constituencies with ELF scores of less than 0.05 because our measure of segregation does not produce meaningful estimates without a minimum level of diversity. Columns 1 and 2 of Tables SI.4 and SI.5 show, however, that our results are not sensitive to including all constituencies in the analysis.

In a second set of robustness tests, we drop all urban constituencies in the sample. Doing so allows us to more effectively control for the demand for boreholes, as the demand for clean water is much lower in urban areas where there is greater access. We code constituencies as urban if the population density is greater than 5000 people per square kilometer. Using this cutoff, we drop 17 constituencies from the sample: Blantyre Bangwe, Blantyre City Central, Blantyre City East, Blantyre City South, Blantyre City South East, Blantyre City West, Blantyre Kabula, Blantyre Malabada, Lilongwe City Central, Lilongwe City South East, Lilongwe City South West, Lilongwe City West, Mulanje Central, Mulanje Limbuli, Mzimba Mzuzu City, Nkhata Bay East, and Zomba Central. Columns 3 and 4 of Tables SI.4 and SI.5 show that our results are robust to the removal of these constituencies.

In a third set of robustness tests, we remove all constituencies in Machinga and Mangochi districts. We do so because these districts were affected by a relatively large rural resettlement program that the government of Malawi established in 2004. The program resettled households from Thyolo and Mulanje districts to Machinga and Mangochi, potentially altering ethnic demographics in the receiving districts. See Chinsinga (2011) for details. Columns 5 and 6 of Tables SI.4 and SI.5 show that our results are robust to the removal of the constituencies in these districts.

Table SI.4: Segregation and Borehole Investments across Constituencies, Different Samples

	Dependent variable:					
	Number of New Boreholes					
	All Constituencies		Rural Only		No Machinga/Mangochi	
	(1)	(2)	(3)	(4)	(5)	(6)
Segregation (continuous)	0.82		1.28		1.38	
	(0.66)		(0.79)		(0.78)	
Dummy for Medium Segregation		0.38		0.42		0.41
		(0.14)		(0.14)		(0.16)
Dummy for High Segregation		0.27		0.39		0.42
		(0.16)		(0.17)		(0.18)
Ethnic Diversity (ELF)	−0.91	−0.85	−0.77	−0.73	−0.79	−0.78
	(0.37)	(0.36)	(0.41)	(0.39)	(0.45)	(0.44)
Population (10,000s)	0.06	0.06	0.07	0.06	0.07	0.07
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Population Density	−0.67	−0.70	0.41	0.18	−0.65	−0.73
	(0.25)	(0.22)	(0.89)	(0.84)	(0.27)	(0.25)
No. Boreholes, 1998	0.05	0.05	0.04	0.04	0.04	0.04
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Intercept†	4.05	4.23	3.60	4.00	3.72	4.13
	(0.44)	(0.37)	(0.51)	(0.42)	(0.51)	(0.42)
Admin. District FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	193	193	166	166	164	164

†Admin. district intercept for Chikwawa displayed

Table SI.5: Segregation and Within-Constituency Targeting of Boreholes, Different Samples

	Dependent variable:					
	I(EA Received Borehole)					
	All Constituencies with >0 new boreholes		Rural Only		No Machinga/Mangochi	
	(1)	(2)	(3)	(4)	(5)	(6)
Ethnic Match	−0.04	−0.0003	0.13	0.10	0.08	0.03
	(0.38)	(0.20)	(0.47)	(0.23)	(0.40)	(0.22)
Ethnic Match x Segregation	0.71		0.52		0.19	
	(0.87)		(1.02)		(0.91)	
Ethnic Match x Med. Segregation		0.64		0.60		0.42
		(0.26)		(0.29)		(0.27)
Ethnic Match x High Segregation		0.09		0.08		−0.09
		(0.29)		(0.31)		(0.30)
Population (1000s)	0.74	0.74	0.83	0.83	0.68	0.68
	(0.08)	(0.08)	(0.10)	(0.10)	(0.08)	(0.08)
Population Density	−0.30	−0.30	0.67	0.66	−0.30	−0.30
	(0.10)	(0.10)	(0.32)	(0.33)	(0.11)	(0.11)
No. Boreholes, 1998	−0.60	−0.60	−0.71	−0.71	−0.62	−0.63
	(0.09)	(0.09)	(0.10)	(0.10)	(0.09)	(0.09)
Intercept†	−1.07	−0.85	−1.31	−1.04	−0.85	−0.59
	(0.15)	(0.23)	(0.22)	(0.22)	(0.22)	(0.22)
Constituency FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
No. Localities (EAs)	9890	9890	7570	7570	8474	8474
No. Constituencies	157	157	148	148	130	130

†Constituency intercept for Chitipa North displayed

Robustness Tests: Conditional Logit and Linear Probability Models

As we note in the paper, a conditional maximum likelihood estimator due to Chamberlain (1980) can be used to consistently estimate the structural parameters in fixed effects logit models. Another approach is to estimate linear probability models with OLS. Table SI.6 shows that our results are robust to using either approach. This is not surprising given that the bias in the unconditional estimator, which we use in the paper, is not severe when the number of observations within each cluster reaches about 10 (Katz 2001; Greene 2012, Ch. 17), which is the case for all but one of our constituencies.

Table SI.6: Within-Constituency Targeting of Boreholes, Conditional Logit and Linear Probability Models

	Dependent variable:			
	I(EA Received Borehole)		I(EA Received Borehole)	
	Conditional Logit		*OLS*	
	(1)	(2)	(3)	(4)
Segregation (continuous)	−0.04	0.001	0.004	−0.0003
	(0.33)	(0.14)	(0.05)	(0.02)
Dummy for Medium Segregation	0.71		0.11	
	(0.74)		(0.12)	
Dummy for High Segregation		0.63		0.13
		(0.20)		(0.04)
Ethnic Diversity (ELF)		0.09		0.02
		(0.22)		(0.04)
Population (10,000s)	0.70	0.70	0.10	0.10
	(0.06)	(0.06)	(0.01)	(0.01)
Population Density	−0.29	−0.29	−0.01	−0.01
	(0.04)	(0.04)	(0.002)	(0.002)
No. Boreholes, 1998	−0.66	−0.66	−0.12	−0.12
	(0.09)	(0.09)	(0.01)	(0.01)
Intercept†			0.31	0.35
			(0.05)	(0.06)
Constituency FE	*Yes*	*Yes*	*Yes*	*Yes*
Observations	9,482	9,482	9,482	9,482

Robustness Tests: Controlling for Other Goods

In this section, we show that our main results are robust to controlling for (1) the broader stock of public goods in each constituency and locality as of 1998 and (2) the provision of public goods other than boreholes between 1998 and 2008. We control for the former because the stock of public goods may be related to an area's relative demand for clean water. For example, localities without a health clinic and without a borehole may exhibit less demand for a new borehole than localities that have a clinic but no borehole. We thus control for the stock of health clinics and schools in each constituency and locality. The second analysis allows us to account for the fact that the provision of certain goods might be substitutes. We thus control for the number of new clinics and schools built between 1998 and 2008. The results presented in Tables SI.7 and SI.8 show that our results are robust to the inclusion of this expanded set of controls.

Table SI.7: Segregation and Borehole Investments across Constituencies, Controlling for Presence of Clinics and Schools

	Dependent variable:					
	Number of New Boreholes					
	(1)	(2)	(3)	(4)	(5)	(6)
Segregation (continuous)	0.90		1.31		0.86	
	(0.71)		(0.74)		(0.69)	
Dummy for Medium Segregation		0.36		0.42		0.29
		(0.14)		(0.15)		(0.14)
Dummy for High Segregation		0.30		0.37		0.22
		(0.16)		(0.17)		(0.16)
Ethnic Diversity (ELF)	−0.71	−0.70	−0.88	−0.76	−0.86	−0.81
	(0.37)	(0.36)	(0.40)	(0.39)	(0.36)	(0.36)
Population (10,000s)	0.01	0.005	0.08	0.07	−0.05	−0.05
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)
Population Density	−0.50	−0.56	−0.79	−0.87	−0.45	−0.52
	(0.23)	(0.21)	(0.26)	(0.23)	(0.23)	(0.21)
No. Boreholes, 1998	0.05	0.04			0.04	0.04
	(0.01)	(0.01)			(0.01)	(0.01)
No. Clinics, 1998	−0.02	−0.01			−0.02	−0.01
	(0.03)	(0.03)			(0.03)	(0.03)
No. Schools, 1998	0.02	0.02			0.04	0.04
	(0.01)	(0.01)			(0.01)	(0.01)
No. New Clinics, 1998-2008			0.05	0.04	0.02	0.02
			(0.04)	(0.04)	(0.04)	(0.04)
No. New Schools, 1998-2008			−0.01	−0.01	0.03	0.03
			(0.01)	(0.01)	(0.01)	(0.01)
Intercept†	2.67	2.86	2.90	3.19	2.63	2.86
	(0.52)	(0.42)	(0.55)	(0.43)	(0.51)	(0.41)
Admin. District FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	183	183	183	183	183	183

†Admin. district intercept for Chikwawa displayed

Table SI.8: Segregation and Within-Constituency Targeting of Boreholes, Controlling for Presence of Clinics and Schools (Binary Ethnic Match)

	Dependent variable:					
	I(EA Received Borehole)					
	(1)	(2)	(3)	(4)	(5)	(6)
Ethnic Match	−0.05	−0.01	−0.05	−0.02	−0.04	−0.02
	(0.39)	(0.20)	(0.39)	(0.20)	(0.40)	(0.20)
Ethnic Match x Segregation	0.75		0.70		0.72	
	(0.88)		(0.90)		(0.90)	
Ethnic Match x Med. Segregation		0.66		0.65		0.68
		(0.26)		(0.27)		(0.26)
Ethnic Match x High Segregation		0.12		0.10		0.12
		(0.29)		(0.30)		(0.29)
Population (1000s)	0.65	0.65	0.68	0.68	0.63	0.63
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Population Density	−0.27	−0.27	−0.28	−0.28	−0.26	−0.26
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
No. Boreholes, 1998	−0.71	−0.71			−0.72	−0.73
	(0.10)	(0.10)			(0.10)	(0.10)
No. Clinics, 1998	0.25	0.25			0.21	0.21
	(0.11)	(0.11)			(0.11)	(0.11)
No. Schools, 1998	0.46	0.46			0.48	0.49
	(0.06)	(0.06)			(0.06)	(0.05)
No. New Clinics, 1998-2008			0.09	0.09	0.06	0.06
			(0.14)	(0.15)	(0.14)	(0.15)
No. New Schools, 1998-2008			0.50	0.50	0.55	0.55
			(0.09)	(0.09)	(0.09)	(0.09)
Intercept†	−1.07	−0.86	−1.14	−0.92	−1.12	−0.90
	(0.15)	(0.22)	(0.16)	(0.23)	(0.15)	(0.23)
Constituency FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
No. Localities (EAs)	9482	9482	9482	9482	9482	9482
No. Constituencies	148	148	148	148	148	148

†Constituency intercept for Chitipa North displayed

Robustness Tests: Ethnic Match By Term

In 2006, a Constituency Development Fund (CDF) was introduced in Malawi. The CDF provides MPs with resources with which to sponsor development projects in their constituencies. In our main within-constituency analyses analyses, we do not leverage the establishment of the CDF. This is because we do not know whether the boreholes we examine were built before or after 2005. However, because the CDF provides MPs with discretionary control over resources that were previously not at their disposal, it is reasonable to expect that a locality's ethnic connection to the MP after 2005 may have a larger effect on the likelihood of receiving a boreholes than the connection in before 2005.

To test this expectation, we generate our dichotomous and continuous ethnic match measures by term. Table SI.9 presents the results. While the patterns in both terms are consistent with our theoretical framework, there appears to be more favoritism toward coethnic localities in medium and high segregation constituencies (but not in low segregation ones) when we code match with the MP after 2004. This is consistent with the notion that the CDF provided MPs with greater access to discretionary resources that MPs targeted in ways predicted by our model. We note, however, that these findings are only suggestive and are limited because we cannot link the construction of boreholes to each specific legislative term.

Table SI.9: Segregation and Within-Constituency Targeting of Boreholes, By Term

	Dependent variable:							
	I(EA Received Borehole)							
	Binary Match				Continuous Match			
	1999	1999	2004	2004	1999	1999	2004	2004
Ethnic Match	0.07	0.13	0.003	−0.06	0.02	0.42	−0.14	0.05
	(0.41)	(0.20)	(0.40)	(0.18)	(0.96)	(0.41)	(0.90)	(0.27)
Match x Seg.	0.24		0.53		1.17		1.58	
	(0.93)		(0.88)		(2.08)		(1.97)	
Match x Med. Seg.		0.26		0.58		0.36		0.86
		(0.29)		(0.26)		(0.55)		(0.42)
Match x High Seg.		−0.21		0.24		−0.02		0.57
		(0.30)		(0.28)		(0.57)		(0.45)
Population (1000s)	0.70	0.70	0.73	0.73	0.70	0.70	0.72	0.72
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Population Density	−0.30	−0.30	−0.30	−0.30	−0.29	−0.29	−0.29	−0.29
	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.10)	(0.10)
Boreholes, 1998	−0.62	−0.62	−0.69	−0.69	−0.62	−0.63	−0.70	−0.70
	(0.12)	(0.12)	(0.09)	(0.09)	(0.12)	(0.12)	(0.09)	(0.09)
Intercept†	−0.75	−0.71	−0.96	−0.89	−0.84	−0.80	−1.20	−1.17
	(0.07)	(0.07)	(0.13)	(0.20)	(0.08)	(0.11)	(0.18)	(0.28)
Constituency FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
No. Localities (EAs)	9530	9530	9495	9495	9411	9411	9495	9495
No. Constituencies	145	145	147	147	144	144	147	147

†Median constituency intercept (Machinga North East) displayed

Other Public Goods: Health Clinics and Schools

In this section, we show that our theory also accounts for the provision of health clinics and schools. This is important because inferences about distributive politics can be sensitive to the particular public good under study (Kramon and Posner 2013). Tables SI.10 and SI.11 show that our results are robust to the public good that we investigate. At the constituency level, the provision of clinics and schools is increasing with the level of segregation of the constituency. At the locality level, there is evidence of ethnic favoritism in the provision of schools and clinics at higher levels of segregation, but none where ethnic groups are integrated. These findings provide greater confidence in our theoretical framework.

The analyses presented in this section also allow us to test an implication of our theory: that the degree of segregation necessary for a good's provision and ethnic targeting is increasing in the *geographic reach* of the good. Since the reach of schools and clinics is typically larger than the reach of boreholes, our framework implies that we should observe greater investment and favoritism in school and clinic provision in the most segregated constituencies and greater investment and favoritism in borehole provision in the moderately segregated constituencies (which we demonstrate in Tables 1 and 2).

Consistent with these expectations, the results in Columns 2 and 4 of Table SI.10 show that the most clinics and schools were produced in the most highly segregated constituencies. In addition, the results in Table SI.11 provide some evidence that ethnic favoritism in the provision of clinics and schools is concentrated in the most highly segregated constituencies.

Table SI.10: Investments in Clinics and Schools across Electoral Constituencies

	Dependent variable:			
	No. New Clinics		No. New Schools	
	(1)	(2)	(3)	(4)
Segregation (continuous)	0.43		1.33	
	(0.90)		(0.70)	
Dummy for Medium Segregation		0.30		0.56
		(0.23)		(0.18)
Dummy for High Segregation		0.49		0.83
		(0.27)		(0.22)
Ethnic Diversity (ELF)	1.45	1.57	0.38	0.53
	(0.54)	(0.52)	(0.42)	(0.41)
Population (10,000s)	0.10	0.10	0.14	0.15
	(0.03)	(0.03)	(0.02)	(0.02)
Population Density	−0.17	−0.14	0.06	0.05
	(0.13)	(0.11)	(0.06)	(0.05)
No. Clinics, 1998	−0.13	−0.13		
	(0.04)	(0.04)		
No. Schools, 1998			−0.05	−0.05
			(0.01)	(0.01)
Intercept†	−0.53	−0.68	0.82	0.92
	(0.75)	(0.56)	(0.58)	(0.46)
Admin. District FE	*Yes*	*Yes*	*Yes*	*Yes*
Observations	183	183	183	183

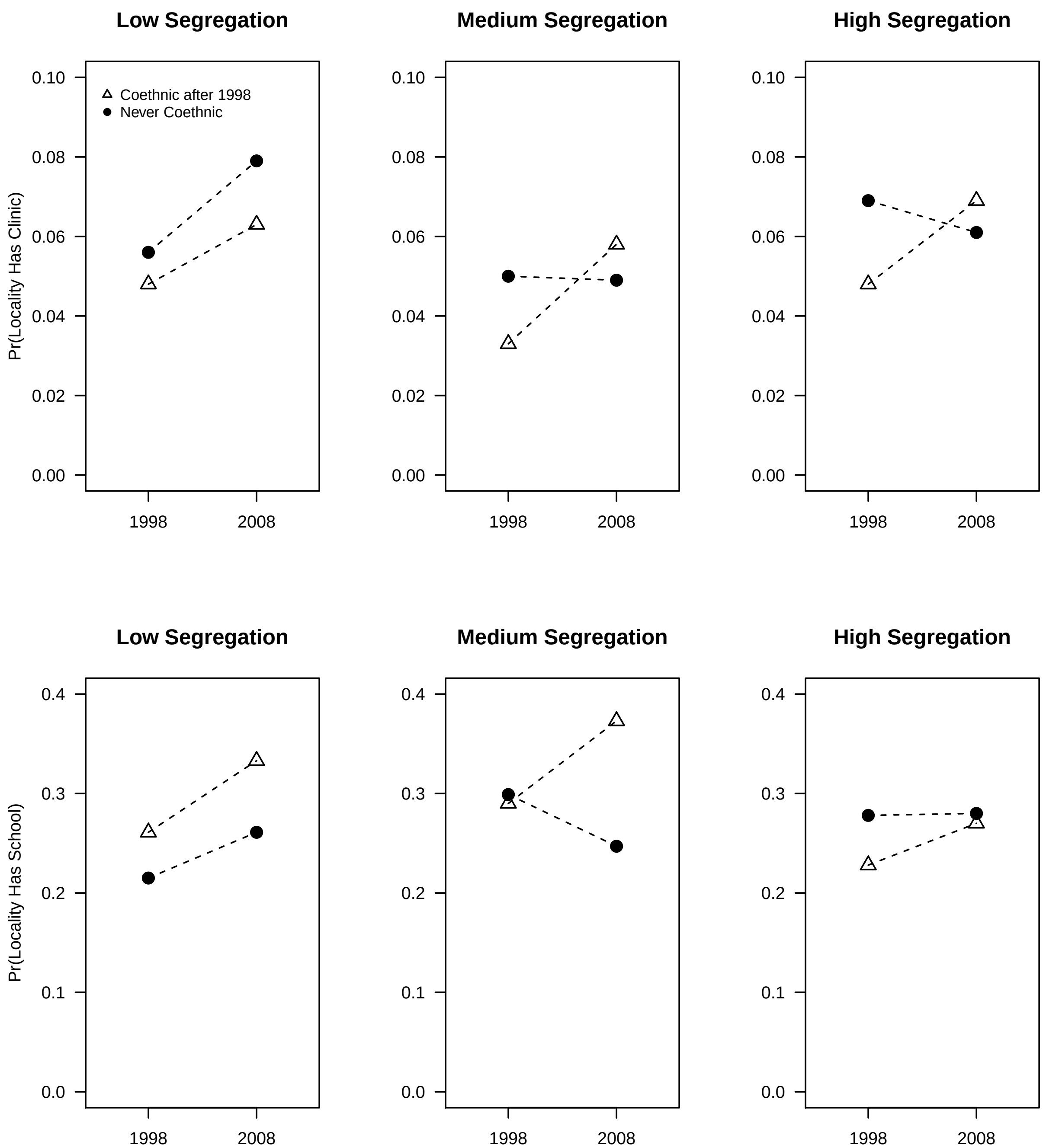
†Admin. district intercept for Chikwawa displayed

Table SI.11: Segregation and Within-Constituency Targeting of Clinics and Schools

	Dependent variable:							
	I(New Clinic)				I(New School)			
	Binary Match		Continuous Match		Binary Match		Continuous Match	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ethnic Match	−1.73	−0.87	−1.43	−1.86	−0.43	−0.08	−0.33	−0.45
	(0.94)	(0.28)	(2.74)	(1.02)	(0.32)	(0.21)	(1.23)	(0.65)
Match x Segregation	4.48		2.19		1.22		0.40	
	(2.16)		(6.04)		(0.77)		(2.69)	
Match x Med. Seg.		1.23		1.81		0.06		0.39
		(0.47)		(1.22)		(0.32)		(0.73)
Match x High Seg.		2.77		1.91		0.46		0.39
		(0.84)		(1.41)		(0.33)		(0.83)
Population (1000s)	0.53	0.52	0.47	0.48	0.40	0.40	0.42	0.42
	(0.24)	(0.24)	(0.20)	(0.20)	(0.12)	(0.12)	(0.11)	(0.11)
Population Density	−0.06	−0.06	−0.09	−0.09	−0.07	−0.07	−0.07	−0.07
	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)
No. Clinics, 1998	−1.16	−1.14	−1.33	−1.32				
	(0.67)	(0.67)	(0.74)	(0.74)				
No. Schools, 1998					−0.74	−0.74	−0.77	−0.77
					(0.17)	(0.17)	(0.17)	(0.17)
Intercept†	−4.11	−5.52	−3.40	−3.60	−2.39	−2.59	−2.17	−2.20
	(0.39)	(0.85)	(0.37)	(0.52)	(0.20)	(0.29)	(0.17)	(0.26)
Constituency FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
No. Localities (EAs)	5735	5735	6039	6039	8954	8954	9231	9231
No. Constituencies	86	86	88	88	138	138	140	140

†Constituency intercept for Machinga East displayed

Figure SI.5: DiDs for Clinics (Upper Panel) and Schools (Lower Panel)



Note: 3502 localities (enumeration areas) located in 120 constituencies are included in the analyses. All of these localities were *not* coethnic with their MP in 1998. 1599 localities became coethnic with their MP in either the 1999 or the 2004 parliamentary elections; these are denoted with a triangle. The 1903 localities denoted with a circle were never coethnic with their MP in the study period.

Heterogeneous Effects by Electoral Competitiveness

It may be that representatives in competitive districts attempt to engage in more ethnic targeting than representatives in "safe" districts. Thus, if local public goods allocations are one way to boost turnout or affect vote choice, the effect of segregation on public goods allocations may be stronger where representatives are less electorally secure.

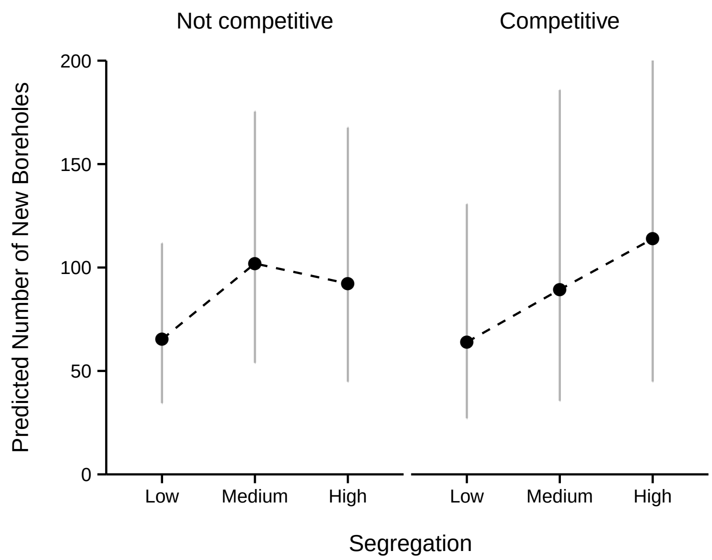
Our data allow only for a limited test of this idea. In particular, we are limited by the fact that we cannot attribute new public goods investments to a given representative or electoral term, as the time period for which we can measure public goods investments (1998-2008) includes two terms. We are also limited by the relatively small number of constituencies with close elections. Analyses of how the effect of segregation on public goods distribution is conditioned by electoral competitiveness therefore remains a fruitful avenue for future research.

Nevertheless, we carry out a suggestive test of this idea by collecting data on the results from the 2004 parliamentary elections.[2] These data are available for 174 constituencies. We code *Winmargin* as the difference between the percentage won by the incumbent-elect and the runner-up in a constituency. Because we have strong reasons to believe that the effect of this variable is non-linear (i.e., there should be little difference between candidates who won by 40 points versus 45 points), we dichotomize constituencies by whether they had a "competitive" election or not. To ensure that the results are not driven by any particular cutoff, we code two such variables based on a win margin of 5 and 10 percentage points. Based on this coding, 21 constituencies are competitive using the 5-point cutoff (i.e., the win margin is less than 5 percentage points in 21 constituencies), and 40 are competitive using the 10-point cutoff.

We find suggestive evidence that segregation matters more for borehole investments in competitive constituencies. The full results are presented in Table SI.12. Based on Model 2 from this table, Figure SI.6 suggests that there is a steeper segregation effect in competitive constituen-

[2] Scraped from `http://www.sdnp.org.mw/election/ele2004/par_results.htm`

Figure SI.6: Predicted new boreholes, by segregation and competitiveness in the 2004 elections



The figure shows the relationship between segregation and borehole investments across competitive and non-competitive constituencies. “Competitive constituencies” are those in which the incumbent won by less than 5 percentage points in the 2004 elections. The expected values (with 95% confidence intervals) are based on 10,000 simulations of Model 2 in Table SI.12.

cies than in non-competitive constituencies. MPs in constituencies with low levels of segregation invested in about the same number of boreholes (65) regardless of whether the 2004 elections were competitive or not, whereas MPs in highly segregated *and* competitive constituencies invested in 20 more boreholes than MPs in highly segregated non-competitive districts, on average. Note, however, that these analyses are under-powered given a small number of competitive districts, resulting in relatively large standard errors. This again highlights a potential opportunity to collect data that are more uniquely suited to test this hypothesis.

We also test whether ethnic favoritism in the distribution of boreholes within constituencies is more pronounced in segregated constituencies with competitive elections. These results are presented in Table SI.13. We split the sample of EAs by whether the constituency in which the EA is located had a competitive election, again employing 5 and 10-point cutoffs to define competitiveness. Using the 5-point cutoff, the effect of segregation and ethnic match appears to be

Table SI.12: Segregation, Electoral Competitiveness, and Borehole Investments across Constituencies

	Dependent variable:			
	Number of New Boreholes			
	Competitive = Win Margin < 5		Competitive = Win Margin < 10	
	(1)	(2)	(3)	(4)
Segregation (continuous)	1.24		1.18	
	(0.82)		(0.84)	
Dummy for Medium Segregation		0.45		0.44
		(0.16)		(0.17)
Dummy for High Segregation		0.34		0.36
		(0.19)		(0.20)
Segregation x Competitive	0.63		0.61	
	(1.94)		(1.56)	
Medium Seg. x Competitive		−0.13		0.05
		(0.46)		(0.37)
High Seg. x Competitive		0.22		0.11
		(0.43)		(0.38)
Competitive	−0.34	−0.05	−0.29	−0.12
	(0.90)	(0.34)	(0.70)	(0.29)
Ethnic Diversity (ELF)	−0.82	−0.76	−0.83	−0.73
	(0.42)	(0.40)	(0.42)	(0.41)
Population (10,000s)	0.07	0.06	0.07	0.06
	(0.02)	(0.02)	(0.02)	(0.02)
Population Density	−0.58	−0.67	−0.57	−0.65
	(0.24)	(0.22)	(0.24)	(0.22)
No. Boreholes, 1998	0.04	0.04	0.04	0.04
	(0.02)	(0.02)	(0.02)	(0.01)
Intercept†	3.83	4.18	3.85	4.16
	(0.51)	(0.41)	(0.52)	(0.41)
Admin. District FE	*Yes*	*Yes*	*Yes*	*Yes*
Observations	165	165	165	165

†Median administrative district intercept (Chikwawa) displayed

stronger in more competitive constituencies (Models 1-2 of Table SI.13) than in less competitive constituencies (Models 3-4). The effect is less clear when we use the 10-point cutoff (see Models 5-8). While these results suggest that it is plausible that electoral competitiveness conditions the relationship between segregation and ethnic favoritism, we again emphasize the need to collect data directly suited for testing this hypothesis. Unfortunately, this is not possible with our current data due to the small number of competitive constituencies and the fact that we cannot know if new boreholes were allocated in the 2004-2009 electoral term (which would be ideal when using 2004 election data) or in the 1999-2004 term.

Table SI.13: Segregation, Electoral Competitiveness, and Borehole Allocations within Constituencies

	Dependent variable:							
	I(EA Received Borehole)							
	Win Margin < 5		Win Margin $\geq$ 5		Win Margin < 10		Win Margin $\geq$ 10	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ethnic Match	−0.14	−0.34	0.02	0.11	0.46	0.05	−0.26	−0.01
	(2.36)	(0.67)	(0.35)	(0.19)	(0.71)	(0.42)	(0.38)	(0.21)
Ethnic Match x Seg.	1.02		0.54		−0.54		1.25	
	(4.90)		(0.81)		(1.66)		(0.82)	
Ethnic Match x Med. Seg.		1.89		0.40		0.72		0.59
		(0.79)		(0.26)		(0.56)		(0.29)
Ethnic Match x High Seg.		0.59		−0.08		−0.29		0.21
		(1.00)		(0.29)		(0.64)		(0.30)
Population (1000s)	0.54	0.59	0.80	0.80	0.58	0.59	0.81	0.80
	(0.24)	(0.23)	(0.09)	(0.09)	(0.14)	(0.14)	(0.10)	(0.10)
Population Density	−0.68	−0.69	−0.28	−0.29	−0.15	−0.15	−0.36	−0.36
	(0.43)	(0.45)	(0.10)	(0.10)	(0.13)	(0.13)	(0.12)	(0.12)
No. Boreholes, 1998	−0.96	−0.95	−0.66	−0.66	−0.67	−0.68	−0.71	−0.71
	(0.31)	(0.31)	(0.10)	(0.10)	(0.22)	(0.22)	(0.11)	(0.11)
Intercept	−0.68	−0.47	0.18	0.20	−0.81	−0.74	0.28	0.31
	(0.31)	(0.36)	(0.12)	(0.13)	(0.15)	(0.21)	(0.13)	(0.14)
Constituency FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
No. Localities (EAs)	1204	1204	7693	7693	1978	1978	6919	6919
No. Constituencies	20	20	116	116	32	32	104	104

Note: All models use *Binary Match* to define the ethnic linkage between MPs and EAs.

Segregation and the Provision of Private Goods

This section examines whether segregation also affects private goods transfers across and within constituencies. These analyses serve primarily as a placebo test for our public goods analyses. In particular, some unobserved characteristic of constituencies may be correlated with both the degree of segregation and the quality of the MP, producing a spurious relationship between ethnic segregation and the quantity of new local public goods. Similarly, some localities may be better able to get a coethnic leader elected *and* more effective in lobbying for new public investments. However, if this were the case, we would expect segregation to be positively associated with investment in, and the ethnic targeting of, *all* types of distributive goods. These analyses thus help us to rule out that selection effects are driving the patterns we observe in the local public goods data. In addition to serving as a placebo test, these analyses provide a limited test of an additional observable implication of our theory. Because incumbents should exert less effort to invest in public goods in less segregated constituencies, they should be more willing to serve their coethnics in other ways — for instance, by providing private goods.

We study the largest and most politically salient form of private transfer from the Malawian government to citizens: coupons given to individual households to subsidize the cost of fertilizer and other agricultural inputs through the Targeted Input Program (TIP) introduced in 2000 (see Harrigan 2008, for an overview of TIP and earlier programs).[3] The vast majority of Malawians are subsistence farmers growing maize for household consumption. In recent years, population pressures and reduced soil quality have resulted in declining productivity and increased food insecurity. In response to declining agricultural productivity and increased food insecurity, the Government of Malawi instituted a number of programs, culminating in the Targeted Input Program (TIP). TIP was

[3] TIP was eventually replaced by a larger scale subsidy program in 2005, after our data were collected.

introduced in 2000 and scaled up after the 2002 famine.[4] The program provided seeds, fertilizer, and other agricultural inputs to households, especially the most vulnerable households. Between 2000 and 2004, an estimated 7 million beneficiaries received inputs through TIP (Harrigan 2008).

While TIP was designed to be programmatic (Chinsinga 2005), in practice political elites exercised considerable discretion over the distribution of subsidies (Chasukwa, Chiweza, and Chikapa-Jamali 2014; Øygard et al. 2003; Tambulasi 2009) and evidence suggests that political elites utilized that discretion politically (Mason and Ricker-Gilbert 2013; Ricker-Gilbert and Jayne 2011).[5]

Ideally, we would analyze data on the total number of subsidies distributed within each constituency, and the geographic location and ethnicity of each recipient within constituencies. Unfortunately, such data do not exist. Instead, we utilize a nationally representative survey data gathered as part of Malawi's second Integrated Household Survey (IHS2), which records whether a household received a TIP coupon during the three years prior to the survey. The survey was designed by the Government of Malawi's National Statistics Office and was implemented with support from the World Bank and the International Food Policy Research Institute (IFPRI). Data were collected on $11,279$ Malawians residing in 560 randomly selected Enumeration Areas between March 2004 and March 2005. Fully 53% of the sample received a TIP transfer.

The sampling procedure of the IHS2 is as follows. The sample includes all three regions: north, center, and south. The country was first stratified into urban and rural strata. Urban areas

[4]TIP was eventually replaced by a larger scale subsidy program in 2005, after our data were collected.

[5]Dionne and Horowitz (2013) find no evidence of ethnic targeting in the distribution of TIP's successor program, Malawi's Agricultural Input Subsidy Program, in three Malawian districts between 2008 and 2009. However, they only evaluate whether coethnics of the president or member of the largest three groups were favored, and do not consider the effect of sharing an ethnicity with one's MP.

include the four major urban centers: Lilongwe, Blantyre, Mzuzu, and the Municipality of Zomba. The rural strata were further broken down into 27 additional strata corresponding to Malawi's 27 administrative districts. One district, Likoma, was excluded because it is an island and difficult to travel to. The sampling was therefore stratified into 30 strata: 26 districts and four urban areas.

In the first stage of the sampling procedure, EAs were randomly selected from within each strata. The number of EAs selected was proportional to the total size of the strata: 12 EAs from those with 0 to 75,000 households; 24 EAs from those with 75,000 to 125,000 households; 36 EAs from those with 125,000 to 175,000 households; and 48 EAs from those with 175,000 to 225,000 households. In the second stage, 20 households were selected at random from within each of the EAs chosen in the first stage. Figure SI.7 maps the EAs for which we have data. Because of the random sampling of EAs, only 148 of the 193 constituencies had sufficient data to include in our analysis.[6]

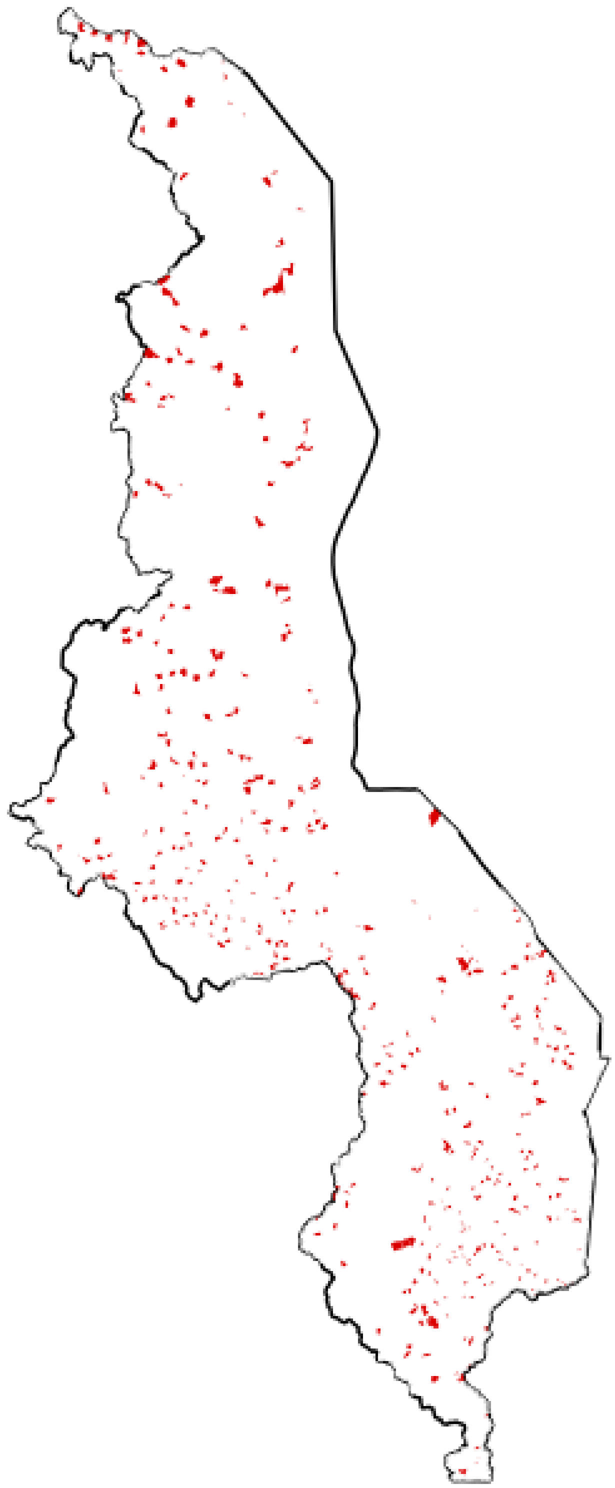
We first analyze the relationship between constituency-level segregation and overall investments in TIP transfers. The dependent variable is the proportion of households in the constituency that received a transfer. Because this measure is based solely on a sample, it is measured with error.[7] We deal with this partially in our analyses by weighting each constituency by the inverse of the standard error of the constituency level estimate (following Saxonhouse 1976). Since the dependent variable at the constituency level is the proportion of households receiving a TIP transfer, we implement a fractional logistic regression model (Papke and Wooldridge 1996). Because TIP was designed to benefit the poor and ultra-poor, we control for the proportion of the sample within

---

[6]The IHS2 sample was generated by random selection of EAs, and (by chance) did not include data from 45 electoral constituencies. Figure SI.7 of the online appendix shows the distribution of sampled EAs.

[7]On average, each constituency sample includes 132 individuals, ranging from 39 to 344 constituents.

Figure SI.7: EAs Included in the Malawi Integrated Household Survey (IHS2) Sample



each constituency that is classified as such by the IHS2, in addition to controls for ethnic diversity, urban center, and population.

Results are presented in Table SI.14. Column 1 presents results with the continuous measure of segregation and no controls. In stark contrast to the public goods results, ethnic segregation has a *negative* and significant relationship with investments in private transfers. We estimate that about 67% of households in an integrated constituency (10th percentile of segregation) received the TIP transfer. In segregated areas (90th percentile), 57% of households received the transfer. In Model 2, which adjusts for confounding, the coefficient on segregation remains negative and large, but is no longer statistically significant. In Model 3 we include the dummy indicators of segregation levels. The coefficient on middle segregation is positive, but very close to zero, while the coefficient on high segregation is negative, but not statistically significant.

We also examine ethnic favoritism in TIP allocations *within constituencies*. To do so, we create a dichotomous ethnic match variable that takes a value of 1 if the household head in the IHS2 survey shares an ethnicity with the MP of the constituency in which the household is located, and 0 otherwise. However, because IHS2 does not ask about ethnicity explicitly, we use language as a proxy for ethnicity.[8] Because the data cover TIP receipts between 2001-2004, we only code individual survey respondents' ethnic linkage to the MP that was elected in their constituency in 1999. We interact this individual-level indicator of ethnic match with constituency-level segregation measures. Since our dependent variable is a binary indicator of receipt of TIP at the individual level, we use a logistic regression, with standard errors clustered by electoral constituency. We include dummies for whether a household is considered poor or ultra-poor, with non-poor households as the omitted category, as well as constituency fixed effects.

---

[8]We code respondent ethnicity as the ethnic group associated with the respondent's home language. In Malawi, language uniquely identifies some but not all ethnic groups. Afrobarometer survey data from 2005, which asks about both ethnic identity and home language, suggest that language is an appropriate indicator of ethnicity for around 75% of the population.

Table SI.14: Segregation and Overall Levels of Private Goods Transfers

	Dependent variable:		
	Proportion Receving Transfer		
	(1)	(2)	(3)
Segregation (continuous)	−1.49	−1.00	
	(0.69)	(0.81)	
I(Medium Segregation)			0.06
			(0.17)
I(High Segregation)			−0.11
			(0.20)
Ethnic Diversity (ELF)		0.20	0.27
		(0.32)	(0.31)
Population (in 10,000s)		−0.03	−0.04
		(0.02)	(0.02)
Includes Urban Center		0.14	0.14
		(0.17)	(0.18)
Poor		0.23	0.16
		(0.99)	(0.96)
Ultra-Poor		0.52	0.63
		(1.31)	(1.28)
Intercept	1.18	0.89	0.45
	(0.33)	(0.56)	(0.39)
Observations	149	149	149

Weighted fractional logistic regression models

Table SI.15 presents the results. Model 1 estimates the extent of ethnic favoritism in all constituencies, and finds that co-ethnics of MPs are about 11 percentage points more likely to receive the TIP transfer than are non-coethnics. Model 2 estimates whether ethnic favoritism is conditioned by a constituency's level of ethnic segregation by including an interaction term. The coefficient on ethnic match, which captures ethnic favoritism in completely integrated constituencies, is positive and quite large. However, consistent with our expectations, the interaction between ethnic match and segregation is negative. While the coefficient on the interaction term is not significant at conventional levels, it is quite large. This is precisely the opposite pattern than we uncover with local public goods, where ethnic favoritism is increasing in segregation. In Model 3, we interact the ethnic match indicator with the two dummy variables for middle and high levels of segregation. These interaction coefficients are very close to zero and are not statistically significant.

In summary, unlike with public goods, private goods provision is not affected by segregation, or may even be decreasing with segregation. The same appears to be true with respect to ethnic favoritism in the transfer of private goods. Thus, the results from this section allay concerns that omitted variable bias is driving the results of our public goods analyses.

Table SI.15: Segregation and Within-Constituency Targeting of Private Goods

	Dependent variable:		
	I(Received Transfer)		
	(1)	(2)	(3)
Ethnic Match (individual)	0.59	0.90	0.61
	(0.14)	(0.44)	(0.19)
Ethnic Match x Segregation		−0.74	
		(1.00)	
Ethnic Match x Med. Segregation			−0.09
			(0.30)
Ethnic Match x High Segregation			0.10
			(0.37)
Poor	0.36	0.36	0.36
	(0.07)	(0.07)	(0.07)
Ultra-Poor	0.40	0.40	0.40
	(0.08)	(0.08)	(0.08)
Intercept	1.19	1.19	1.19
	(0.03)	(0.03)	(0.03)
Constituency FE	Yes	Yes	Yes
No. Individuals	8563	8563	8563
No. Constituencies	143	143	143

References: Supporting Information

Chamberlain, Gary. 1980. "Analysis of Covariance with Qualitative Data." *The Review of Economic Studies* 47: 225–238.

Chasukwa, Michael, Asiyati L Chiweza, and Mercy Chikapa-Jamali. 2014. "Public Participation in Local Councils in Malawi in the Absence of Local Elected Representatives-Political Eliticism or Pluralism?" *Journal of Asian and African Studies* 49(6): 705–720.

Chinsinga, Blessings. 2005. "District Assemblies in a Fix: The Perils of Self-Seeking Tendencies in Decentralisation Policy Reforms in Malawi." *Development Southern Africa* 22(4): 529–548.

Chinsinga, Blessings. 2011. "The politics of land reforms in Malawi: The case of the Community Based Rural Land Development Programme (CBRLDP)." *Journal of International Development* 23(3): 380–393.

Cutler, David M., Glaeser Edward L., and Jacob L. Vigdor. 2012. "The Rise and Decline of the American Ghetto." *Journal of Political Economy* 107(3): 455–506.

Dionne, Kim Yi, and Jeremy Horowitz. 2013. "The Political Effects of Anti-Poverty Initiatives: An Analysis of Malawi's Agricultural Input Subsidy Program.".

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.

Government of Malawi. 1994. "1994 Presidential and Parliamentary General Election Results." *The Malawi Government Gazette Extraordinary* XXXI(40): 267–282.

Government of Malawi. 1999. "Parliamentary General Election Results." Available from the Sustainable Development Network Programme website at `http://www.sdnp.org.mw`.

Government of Malawi. 2004. "Parliamentary General Election Results." Available from the Sustainable Development Network Programme website at http://www.sdnp.org.mw.

Greene, William H. 2012. *Econometric Analysis*. La Jolla, CA: Granite Hill Publishers.

Harrigan, Jane. 2008. "Food Insecurity, Poverty and the Malawian Starter Pack: Fresh Start or False Start?" *Food Policy* 33(3): 237–249.

Katz, Ethan. 2001. "Bias in Conditional and Unconditional Fixed Effects Logit Estimation." *Political Analysis* 9(4): 379–384.

Kramon, Eric, and Daniel N. Posner. 2013. "Who Benefits from Distributive Politics? How the Outcomes One Studies Affect the Answers One Gets." *Perspectives on Politics* 11(2).

Mason, Nicole, and Jacob Ricker-Gilbert. 2013. "Disrupting Demand for Commercial Seed: Input Subsidies in Malawi and Zambia." *World Development* 45: 75–91.

Matuszeski, Janina, and Frank Schneider. 2006. "Patterns of Ethnic Group Segregation and Civil Conflict." Working Paper, Harvard University. Available at: http://www.cgdev.org/doc/events/02.09.07/Matuszeski-JMP.pdf.

Øygard, Ragnar, R. Garcia, A. Guttormsen, R. Kachule, A. Mwanaumo, I. Mwanawina, E. Sjaastad, M. Wik, A. Bonnerjee, J. Braithwaite, S. Carvalho, K. Ezemenari, C. Graham, and A. Thomson. 2003. "The Maze of Maize: Improving Input and Output Market Access for Poor Smallholders in Southern African Region." Department of Economics and Resource Management, Agricultural University of Norway Report No. 26.

Papke, Leslie E., and Jeffrey M. Wooldridge. 1996. "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates." *Journal of Applied Econometrics* 11: 619–32.

Reardon, Sean F, and David O'Sullivan. 2004. "Measures of Spatial Segregation." *Sociological Methodology* 34(1): 121–162.

Ricker-Gilbert, Jacob, and Thomas S. Jayne. 2011. "What are the Enduring Effects of Fertilizer Subsidy Programs on Recipient Farm Households? Evidence from Malawi." Michigan State University's Department of Agricultural, Food, and Resource Economics Staff Paper No. 2011-09.

Saxonhouse, Gary R. 1976. "Estimated Parameters as Dependent Variables." *American Economic Review* 66(1): 178–183.

Tambulasi, Richard IC. 2009. "The Public Sector Corruption and Organized Crime Nexus: The Case of the Fertiliser Subsidy Programme in Malawi." *African Security Review* 18(4): 19–31.