

# Public Goods, and Nested Subnational Units: Diversity, Segregation, or Hierarchy?

Naveen Bharathi<sup>a</sup>, Deepak Malghan<sup>a</sup>, Sumit Mishra<sup>b</sup>, Andaleeb Rahman<sup>c</sup>

<sup>a</sup>*Indian Institute of Management Bangalore, Bengaluru 560 076, INDIA*

<sup>b</sup>*Institute for Financial Management and Research, Chennai, 600 034, INDIA*

<sup>c</sup>*Cornell University, Ithaca NY 14853, USA*

---

## Abstract

We develop a general multi-scale diversity framework to account for spatial segregation of hierarchically-ordered ethnic groups residing in politically and administratively nested geographic aggregations. We explicate how ethnic diversity, ethnic segregation, and ethnic hierarchy interact with the “public goods catchment area” to cast doubt on extant hypotheses linking diversity and public goods provisioning. We not only show how the celebrated “diversity debit” relationship is incomplete at best but also call into question the more recent literature that posits a positive association between ethnic segregation and public goods. We test our framework using a large national census dataset containing ethnicity information (aggregate caste categories) for  $\approx 830$  million rural residents in India. Our nested-geography models use data from villages ( $n \approx 600,000$ ) and sub-districts containing these villages ( $n \approx 6,000$ ) for twenty-five different public goods. We show how not accounting for the spatial structure of diversity, segregation, and hierarchy biases empirical models of diversity and public goods. To the best of our knowledge, the empirical evidence in this paper comes from the largest dataset used in the politics of public goods literature.

*Key words:* Ethnic Diversity, Ethnic Hierarchy, Nested Spatial Scales, Public Goods Catchment Area

---

---

*Email addresses:* [naveen.bharathi@iimb.ac.in](mailto:naveen.bharathi@iimb.ac.in) (Naveen Bharathi), [dmalghan@iimb.ac.in](mailto:dmalghan@iimb.ac.in) (Deepak Malghan), [sumit.mishra@ifmr.ac.in](mailto:sumit.mishra@ifmr.ac.in) (Sumit Mishra), [ar687@cornell.edu](mailto:ar687@cornell.edu) (Andaleeb Rahman)

2018-09-04 15:38 IST

## Introduction

Empirical political economy scholarship has made a strong case for the negative effect of ethnic diversity on public good provisioning (Alesina et al., 1999). This negative relationship is driven by varying preferences and coordination failures, if not outright strife between ethnic groups (Habyarimana et al., 2007). Recent empirical evidence, however, suggests that the “diversity-debit hypothesis” breaks down at subnational scales (Gerring et al., 2015; Gisselquist et al., 2016). A sense of “we-ness” and a shared regional identity potentially supersedes ethnic divisions and ameliorates negative effects of diversity (Singh, 2015, 2106). Yet, this theory fails to predict disparities in public goods incidence within sub-national boundaries. Attempts have been made to explain local (non)effects of ethnic diversity by highlighting the potential confounding factors: segregation (spatial distribution of ethnic groups), presence of minority dominated regions and ethnic favoritism (Kustov and Pardelli, 2018; Tajima et al., 2018; Ejdemyr et al., 2017).

We contribute to the emerging debates in the politics of diversity and public goods provisioning literature by showing how ethnic diversity, ethnic segregation, and ethnic hierarchy are three analytically distinct but mutually reinforcing channels influencing the ethnic politics of diversity and public goods. This mutual reinforcement of diversity, segregation, and hierarchy is particularly salient when multiple hierarchically nested administrative and political units interact to provide public goods. For example, local village-level public goods are not only impacted by village-level ethnic diversity, but also by patterns of diversity and segregation within the sub-district that contains the village, the district that contains the sub-district, and finally the sub-national province containing the district. In such multi-scale diversity-contexts, local public goods provisioning are determined by how political elites – responsible for the local placement of public goods – interact with various levels of public administration (Rugh and Trounstein, 2011). Political elites can resort to ethnic favoritism and strategically place public goods in segregated homogeneous localities inhabited by their co-ethnics to maximize potential electoral benefits (Ejdemyr et al., 2017). We argue that

ethnic segregation is at best only a necessary condition for ethnic favoritism. Observed patterns of discrimination is also contingent on the spatial structure of ethnic hierarchy that determines the capacity of political actors to actually discriminate in a segregated setting. We empirically validate this conjecture using census data containing ethnic information for  $\approx 830$  million residents of rural India and a comprehensive vector of 25 different public goods from  $\approx 600,000$  villages across India that are nested within  $\approx 6,000$  sub-districts.<sup>1</sup>

### **Multi-scale Diversity Context, Segregation, and Hierarchy**

The spatial structure of ethnic distribution – or ethnic segregation – has not received adequate attention in the political economy of public goods literature despite evidence that segregation can potentially reduce or even reverse the impact of diversity. Collective action problems are more easily solved in segregated spaces. Locally homogeneous neighborhoods also provide an incentive for more diverse ones to engage in “sibling rivalry like” behavior and effectively lobby for public goods (Tajima et al., 2018). With the same level of ethnic diversity, public goods could therefore be higher in more segregated regions. However, not all segregated spaces are created equal. In addition to segregation we investigate how (if) the differing political influence of segregated ethnic groups impact public goods provisioning. Is the ability to lobby and organize demand by ethnic groups impacted by spatial patterns of segregation as well as the relative positions of segregated groups on the social hierarchy? In a decentralized political and administrative structure, advocacy initiatives could be worthless if the more powerful group are able to discriminate against co-ethnics (Lee, 2017). What are the implications of the combined effects of ethnic segregation and ethnic hierarchy on the strength and direction of the linkage between public goods provisioning and diversity? We answer this question by developing a multi-scale diversity framework that jointly accounts for diversity, segregation, and hierarchy.

---

<sup>1</sup>Villages are the lowest geographical units in India, where public good placements are determined in a decentralized manner.

In figure 1, we show why a multi-scale diversity context must account for segregation as much as intra-unit diversity. Consider four hypothetical top-level geographic aggregations,  $Z = \{A, B, C, D\}$ . Each top-level spatial unit is further divided into four sub-units each,  $z = \{a_1 \dots a_4, b_1 \dots b_4, c_1 \dots c_4, d_1 \dots d_4\}$ . Each unit in  $Z$  has identical ethnic shares (represented by colors, black and white), and thus, also identical fractionalization —  $FRA(A) \equiv FRA(B) \equiv FRA(C) \equiv FRA(D)$ . However, the spatial configuration (ethnic segregation) is different in each of these units so that  $SEG(A) < SEG(B) < SEG(C) < SEG(D)$ . Now consider a public good  $P$  whose provision is determined by political economy processes at both spatial scales ( $Z$  and  $z$ ), and has a ‘spatial catchment area’,  $A$  (represented by the circles in the last panel of the figure). If political elites in  $Z$  favour co-ethnics, higher segregation allows for better ethnic targeting of public goods in  $z$ . The catchment area is a function of the particular public good in question – for example, an elementary school will likely have a smaller catchment than a high school.

For public goods provisioning in nested geographies, accounting for segregation is important, but a complete picture of diversity must also account for hierarchies between ethnic groups. In Figure 2, the two larger spatial units ( $A, B \in Z$ ) have identical fractionalization and identical segregation –  $FRA(A) \equiv FRA(B)$ , and  $SEG(A) \equiv SEG(B)$ . Furthermore, each of the smaller constituent units in  $A$  and  $B$ ,  $z = \{a_1 \dots a_4, b_1 \dots b_4\}$ , have the same fractionalization (in our example, the sub-units are perfectly homogeneous with  $FRA(z_i \in A) \equiv FRA(z_j \in B) \equiv 0, \forall i, j$ ). In  $A$ , we have one black sub-unit in a larger white-dominated space; and  $B$  represents a mirror-image with a single white sub-unit in a black-dominated space. How will (if) public good provisioning be different in units  $A$  and  $B$  of Figure 2? We hypothesize that contemporary ethnic stratification, historical ethnic persecution and other drivers of ethnic hierarchy have an impact on the political economy of diversity and public goods. In other words, not all segregated spaces are created equal. For example, a segregated white neighbourhood within a black-dominated city is likely going to experience different outcomes from an isolated black community within a white city.

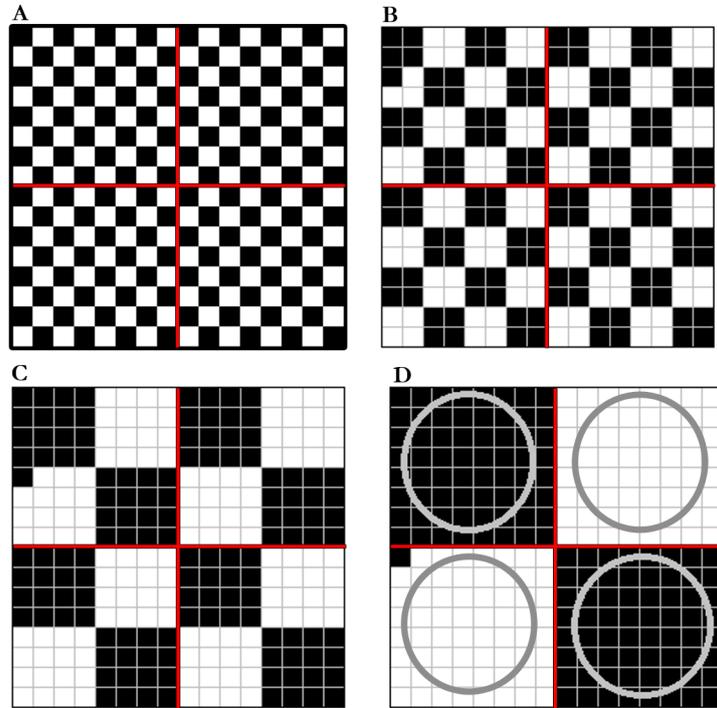


Figure 1: *Segregation, Catchment Area, and Multi-scale Diversity Context*

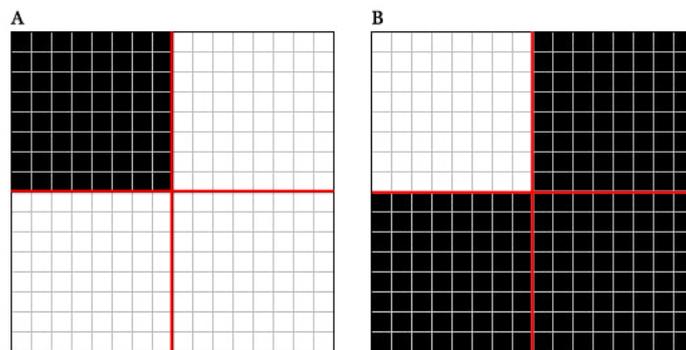


Figure 2: *Ethnic Hierarchy and Nested Geographies*

We formally represent the multi-scale diversity-context relevant for provisioning of  $P$  at spatial-scale  $z$  that is nested in  $Z$  as:

$$D(z \in Z)|_P : \begin{cases} FRA(z), & \text{Intra-unit Fractionalization} \\ SEG(Z), & \text{Inter-unit Segregation} \\ HIE(z, Z), & \text{Ethnic Hierarchy} \\ A(z)|_P, & \text{Relevant Catchment Area} \end{cases} \quad (1)$$

Extant literature has neglected inter-unit segregation, ethnic hierarchy, and catchment area as factors relevant in determining the diversity-development relationship. It is assumed that public goods are provisioned in a “top down” manner (Banerjee et al., 2007) where a single higher-tier of the state is charged with coordination. Even when public goods related policy making might primarily occur at higher echelons of the state, lower-level state actors (especially in in developing country contexts) have considerable discretion to indulge in ethnicity-based discrimination. Budget constraints as well as purposive ethnic coalition building account for such favouritism towards co-ethnics. While geographic segregation abets both these channels, the effects of ethnic hierarchy are are ambiguous, and contingent on the relative ethnic compositions at different levels ( $Z$  and  $z$  in in Equation 1). When ethnic diversity is measured at large geographic aggregates – as is the norm in extant literature – politically salient micro-ecologies of local segregation or ethnic hierarchy in nested political units are overlooked. However, this oversight that is partly driven by data availability, obscures the degree of ethnic tension (or lack of ethnic cooperation) that is at the very heart of the diversity-debit hypothesis (Ejdemyr et al., 2017).

Spatial segregation and ethnic hierarchy are particularly important in nested political contexts that we have described here – homogeneous neighbourhoods (villages) can exist within diverse higher-order geographic aggregations (cities, districts, counties, etc.) that are highly sensitive to ethnic conflicts (Bleaney and Dimico, 2017). In a spatially segregated context, the political implications of ethnic diversity are not easily generalizable – espe-

cially when ethnic groups are also hierarchically stratified. In a decentralized system such as the one under consideration here, while local elections could be thought of as deepening of democracy, they also provide an opportunity for voters to elect “one of their own” in societies where ethnic group markers are salient. This becomes further pronounced especially when such elected representatives have discretion over geographic placement of public goods. When political decentralization is partial and incomplete, multi-aggregation politics is salient as higher administrative aggregations retain substantive residuary powers to influence local-level provisioning of public goods.

The case for studying the multi-scale diversity context rather than simple intra-unit diversity is also bolstered by the collective action and divergent preferences arguments in the empirical literature. Segregation can lead to higher political polarization that inhibits collective action needed to effectively generate demand for public goods that require cooperation across ethnic boundaries (Trounstine, 2016). However, it is easier for homogeneous geographic units to organize and politically articulate their localized public goods demand, especially when ethnic-mixing is low and the state-actor is amenable to demands of a co-ethnic. Discrimination against a homogeneous locality can trigger the mobilization of a rival group demanding similar levels of public goods as the *other*. This “sibling rivalry-like” effects can further contribute to the increase of overall public goods in the segregated region despite higher ethnic diversity (Tajima et al., 2018).

There is also evidence showing how ethnic segregation can increase the incidence of “ethnic voting” (Alesina and Zhuravskaya, 2011). Public goods provisioning can suffer under ethnic voting as electoral outcomes are potentially unhinged from a candidate’s innate ability to deliver for her constituents.<sup>2</sup> The overall effect of regional segregation on local public goods incidence, therefore, remains ambiguous and we know little about how (if) these effects are moderated or amplified when stratified ethnic hierarchy is salient.

---

<sup>2</sup>Cf. Banerjee and Pande (2008) for a detailed exposition on ethnic voting and quality of public services in India – the empirical context of this paper.

## Evidence from India

In order to empirically investigate the relationship between diversity, segregation, and hierarchy across nested spatial units, we use a large national census dataset from India containing ethnicity information in the form of aggregate caste categories for  $\approx 830$  million rural residents of India. Our nested-geography empirical models use data from villages ( $n \approx 600,000$ ) and sub-districts containing these villages ( $n \approx 6000$ ) for twenty-five different public goods with varying catchment areas.

Indian state follows a three-tier federal structure. The governments at the federal level (known as the central, or union government), state level, and local level have direct and indirect roles in public goods provisioning. Typically, allocation of development expenditure including public goods provisioning is made by both federal and state governments which then percolate to districts, sub-districts, and finally to the lowest tier of government – the *panchayats* which are the lowest tier of representative governments and represent small clusters of villages.<sup>3</sup> In such a partially decentralized political and administrative structure, candidates running for elected offices can narrowly appeal a core ethnic constituency, and targeted placement of public services is a much-used channel. For example, ethnic favouritism has been documented in allocation of public goods by the elected head of the *panchayat* benefiting *panchayat*-headquarter village or the head’s co-ethnics (Besley et al., 2004, 2007).

Caste boundaries represent the most important ethnic cleavage in rural India. Steeped in historical notions of ritual purity associated with traditional agrarian occupations, there is considerable overlap between caste and social-human development including income, wealth, and educational attainment (Zacharias and Vakulabharanam, 2011). Caste is a significant

---

<sup>3</sup>The 595,906 villages in our dataset are clustered into 238,617 *panchayats* for an average of  $\approx 2.5$  villages per *panchayat*.

barrier that impedes collective action and its role in determining public goods provisioning is well-established (Banerjee and Somanathan, 2007). Varying group preferences, if not conflicts, punctuate both the location and nature of public goods in segregated Indian villages (Munshi and Rosenzweig, 2016). Networks of political connections across nested administrative levels enables powerful caste groups to indulge in ethnic favoritism and better target public goods relative to less connected groups (Lee, 2017).

In their seminal work on public goods in rural India, Banerjee and Somanathan (2007) find a negative association between share of villages with public goods like schools, health centers, electricity, etc., and caste diversity at the federal parliamentary constituency level. Indian parliamentary constituencies are large geographic aggregates that can contain over a thousand villages. Empirical models of diversity and public goods provisioning are sensitive to spatial scales (Gerring et al., 2015), and can potentially pose problems of statistical inference related to the Modifiable Ethnic Unit Problem (MAUP) even if one were to ignore problems arising out of the neglect of how ethnic diversity interacts with ethnic segregation and ethnic hierarchy. In order to overcome these theoretical and empirical limitations, we use data from the most elementary administrative unit – the village – as our principal spatial unit of analysis.

### *Data and Diversity Metrics*

We use the 2011 Indian national census data<sup>4</sup>. The census data contains caste information in the form of three politically salient aggregate categories – Scheduled Castes (SC); Scheduled Tribes (ST); and Others (OTH) – for all Indian villages ( $n = 595,906$ ) which we aggregate into sub-districts ( $n = 5878$ ) to construct a segregation index. The national census (village directory data) also contains incidence information for a diverse array of public goods including education, health, infrastructure, and transport etc. which we use as our

---

<sup>4</sup>This is the latest census data that is available for India. The next decennial census enumeration will be conducted in 2021

dependent variables.<sup>5</sup>

We construct the fractionalization metric (ELF) for all villages – the workhorse metric used in diversity-development literature and represents the probability that two randomly chosen individuals belong to distinct social groups. Using the notation introduced in Equation 1, for any village  $i \in z$ , fractionalization index is simply:

$$FRA_i = 1 - \left( \sum_{\forall k \in i} \pi_{ik}^2 \right) \quad (2)$$

where  $\pi_{ik}$  is the population share of subgroup  $k \in \{SC, ST, OTH\}$  in village  $i$ . The villages are contained within sub-districts, and we compute the spatial segregation for sub-district  $j \in Z$  using the method of [Goodman and Kruskal \(1954\)](#):

$$SEG_j = \sum_{\forall i \in j} \frac{n_i}{n_j} \left( 1 - \frac{FRA_i}{FRA_j} \right) \quad (3)$$

where  $n_i$  and  $n_j$  are populations of village  $i$  and sub-district  $j$  respectively.  $SEG_j$  represents the extent to which the fractionalization indices computed at the village-level for all villages in sub-district  $j$  are different from the sub-district level fractionalization index.

We denote ethnic hierarchy in a nested-geography space by comparing ethnic dominance – calculated as group with the largest population shares. Let  $D(Z)$  be the dominant group in  $Z$ , the aggregate spatial level, and  $D(z)$  be the dominant group in a sub-unit of  $Z$ . Ethnic hierarchy ( $HIE_{ij}$ ) for a geographic unit  $i \in z$  nested within a larger unit  $j \in Z$  is then a simple ordered pair derived as a Cartesian product of  $D(z)$  and  $D(Z)$ :

$$HIE_{ij} = D(z) \times D(Z); i \in z, j \in Z \quad (4)$$

In our empirical context, we combine the two marginalized groups (SC and ST) into a com-

---

<sup>5</sup>Refer to Table 4 for a full list of our dependent variables.

pound “SCST” category so that we have two groups (SCST and OTH) resulting in four different levels for  $HIE_{ij}$  in Equation 4.

The spatial distribution of sub-district segregation is shown in the left panel (Panel B) of Figure 3. The panel shows segregation quartiles. In Panel B, sub-districts dominated by the two groups (SCST or OTH) is shown. We can see there is some geographic clustering, however there is no clear correlation between sub-district level segregation or group hierarchy.<sup>6</sup>

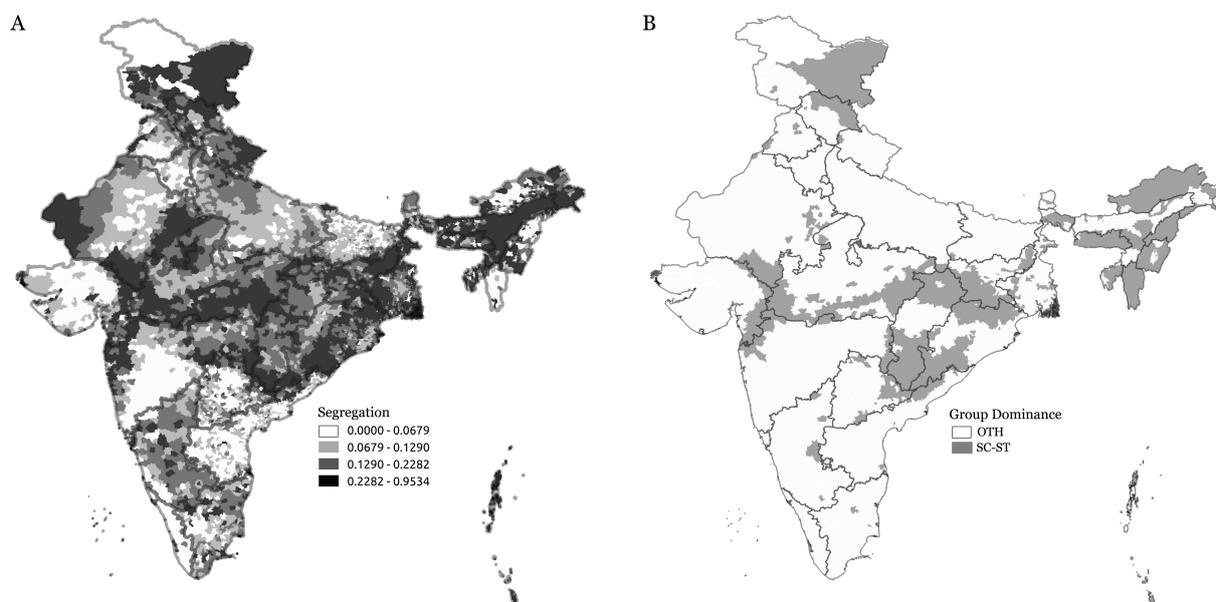


Figure 3: *Sub-district Segregation and Ethnic Dominance.*  $n = 5878$ .

## Results

In order to investigate the joint effects of ethnic diversity, ethnic segregation, and ethnic hierarchy on public goods provisioning, we estimate a village-level linear probability model

<sup>6</sup>In the appendix (Figure 5), we show that there is very little correlation between fractionalization and segregation at the sub-district level.

(LPM) of the following form:

$$Y_{ijp} = \alpha + \underbrace{\beta \times FRA_i}_{\text{Diversity}} + \underbrace{\gamma \times SEG_j}_{\text{Segregation}} + \underbrace{\lambda \times HIE_{ij}}_{\text{Hierarchy}} + \underbrace{\vec{\theta} \circ \vec{V}_i + \vec{\delta} \circ \vec{T}_j}_{\text{Controls}} + \epsilon_i \quad (5)$$

where  $Y_{ijp}$  is the incidence of public good  $p$  in village  $i$  (contained within sub-district  $j$ ). In addition to the village level  $FRA_i$ , we have also included the segregation,  $SEG_j$ , and hierarchy,  $HIE_{ij}$  as variables of interest. Besides district level fixed effects, we include village-level ( $\vec{V}_i$ ), and sub-district level ( $\vec{T}_j$ ) controls vectors.

In the first set of regressions, we include only  $FRA_i$  as the ethnic variable, and the results are presented in Table 3. We find a positive effect of caste diversity on most of the public goods except for health facilities. We progressively include  $SEG_j$  in the next set of models. Sub-district level segregation is found to have the opposite negative effect on the village level public good provisioning (Table ). Statistical significance and sign of the coefficients on  $FRA_i$ , however, are unaffected by the inclusion of  $SEG_j$ . Our findings are contrary to those of [Tajima et al. \(2018\)](#) in two ways. Here, segregation has a negative effect on public goods provisioning, opposite to their findings. Further, we do not find segregation to reduce the effect of diversity. Our results however are in perfect harmony with [Tajima et al. \(2018\)](#) on the following ground – segregation is a key ethnic variable, which is ignored in the diversity-development literature.

Upon the inclusion of ethnic hierarchy variables, along with the segregation and fractionalization, we find all of these ethnic variables to be statistically significant (Table 4). Coefficients on  $FRA_i$ ,  $SEG_j$ , and  $HIE_{ij}$  from this regression are presented in Table 4. There is a clear evidence for diversity-dividend at village-level for most public goods except for the provision of secondary schools and health facilities. Largely, segregation has a negative effect on public goods provisioning. On the effects of village level group hierarchy: villages where the upper castes,  $OTH$ , groups are predominant, incidence of public goods is higher

regardless of ethnic group which are dominant in their sub-district. On the other hand, villages where marginalized group *SCST* are higher in number, public good incidence is lower. Regardless of ethnic diversity or segregation (i.e. despite controlling for  $FRA_i$  and  $SEG_j$ ), ethnic hierarchy seems to be the operative channel which re-inforces public goods incidence.

#### *Effect of diversity by Segregation Quartiles*

Is the association between diversity and public goods mediated by segregation levels? To test for that, we divide our village data into four sub-samples corresponding to the sub-district segregation quartiles (Figure 3), and estimate the same set of models reported in for each of the four sub-samples (Table ). Villages in sub-districts with relatively lower levels of segregation exhibit diversity-debit for a large set of public goods. However, in more segregated sub-districts, village level caste diversity is positively associated with public goods. The logical question ensues: Is there a difference in how group divisions operate to provide public goods depending upon which ethnic groups dominate the segregated regions? We consider sub-districts in the most segregated quartiles where the marginalized groups dominate would have lower public goods.

#### *Ethnic hierarchy in the most segregation sub-district*

Measures of ethnic diversity or segregation are “demographic” measures which are inured to ethnic tensions which reify diversity debit (Lieberman and Singh, 2012). In segregated spaces, it becomes critical to know which groups dominate in order to study any potential discrimination or favoritism. We choose the most segregated quartile and assign a dummy for ethnic hierarchy to study the effect of fractionalization on village level public goods. We find that in the *SCST* villages, diversity has a positive association with public goods if the subdistrict has *OTH* majority (Table 2). On the other hand, we do not any significant effect of diversity in villages where the upper castes are dominant and the lower castes are

dominant at the sub-district level. The association between diversity and public goods, is sensitive to the levels of segregation as well as group hierarchy. Our findings underscore our theoretical claim that diversity, segregation and hierarchy are mutually reinforcing themselves in public good provision in nested administrative structures.

## **Conclusion**

We have theoretically and empirically demonstrated the need to look at multi-scale diversity context rather than simple intra-unit diversity. Using a comprehensive census data set with nearly 600,000 villages, we illustrate how potential biases can arise from the neglect of ethnic segregation across space when public good administration is nested within multiple geographies. Our findings underscore the centrality of spatial unit of analysis in understanding the relationship between diversity and development. Further, our empirical analysis points out to the existing theoretical gaps in understanding how segregation affects the diversity-development association.

## **Acknowledgements**

This paper is expanded from our earlier formulation of segregation (reported in a working paper titled *Spatial Segregation, Multi-scale Diversity, and Public Goods*). This paper is part of a larger project on measuring diversity, segregation, and development. We are grateful to Sanjoy Chakravorty, James Fenske, Ravi Kanbur, Anirudh Krishna, Sharun Mukand, Imran Rasul, Matthias Vom Hau, and Andreas Wimmer. We have used their comments on the larger project in this paper. We alone are responsible for any remaining errors and omissions.

## References

- Alesina, A., Baqir, R., Easterly, W., 1999. Public Goods and Ethnic Divisions. *The Quarterly Journal of Economics* 114 (4), 1243–1284.
- Alesina, A., Zhuravskaya, E., 2011. Segregation and the Quality of Government in a Cross Section of Countries. *The American Economic Review* 101 (5), 1872–1911.
- Banerjee, A., Iyer, L., Somanathan, R., 2007. Public Action for Public Goods. In: *Handbook of Development Economics*. Vol. 4. pp. 3117–3154.
- Banerjee, A., Somanathan, R., 2007. The political economy of public goods: Some evidence from India. *Journal of Development Economics* 82 (2), 287–314.
- Banerjee, A. V., Pande, R., 2008. Parochial Politics: Ethnic Preferences and Politician Corruption.
- Besley, T., Pande, R., Rahman, L., Rao, V., 2004. The politics of public good provision: Evidence from Indian local governments. *Journal of the European Economic Association* 2 (2-3), 416–426.
- Besley, T., Pande, R., Rao, V., 2007. Political Economy of Panchayats in South India. *Economic and Political Weekly* 42, 661–666.
- Bleaney, M., Dimico, A., 2017. Ethnic diversity and conflict. *Journal of Institutional Economics* 13 (02), 357–378.
- Ejdemyr, S., Kramon, E., Robinson, A. L., sep 2017. Segregation, Ethnic Favoritism, and the Strategic Targeting of Local Public Goods. *Comparative Political Studies*, 1–33.
- Gerring, J., Thacker, S. C., Lu, Y., Huang, W., 2015. Does diversity impair human development? A multi-level test of the diversity debit hypothesis. *World Development* 66, 166–188.
- Gisselquist, R. M., Leiderer, S., Niño Zarazúa, M., 2016. Ethnic Heterogeneity and Public Goods Provision in Zambia: Evidence of a Subnational “Diversity Dividend”. *World Development* 78, 308–323.
- Goodman, L. A., Kruskal, W. H., 1954. Measures of association for cross classifications. *Journal of the American statistical association* 49 (268), 732–764.
- Habyarimana, J., Humphreys, M., Posner, D. N., Weinstein, J. M., 2007. Why Does Ethnic Diversity Undermine Public Goods Provision? *The American Political Science Review* 101 (4), 709–725.
- Kustov, A., Pardelli, G., 2018. Ethnoracial homogeneity and public outcomes: The (non) effects of diversity. *American Political Science Review*, 1–8.
- Lee, A., 2017. Ethnic Diversity and Ethnic Discrimination: Explaining Local Public Goods Provision. *Comparative Political Studies* 10 (51), 1351–1383.
- Lieberman, E. S., Singh, P., 2012. Conceptualizing and Measuring Ethnic Politics: An Institutional Complement to Demographic, Behavioral, and Cognitive Approaches. *Studies in Comparative International Development* 47 (3), 255–286.
- Munshi, K., Rosenzweig, M., 2016. Insiders and Outsiders: Local Ethnic Politics and Public Goods Provision, Working Paper.
- Rugh, J. S., Trounstine, J., 2011. The Provision of Local Public Goods in Diverse Communities: Analyzing Municipal Bond Elections. *The Journal of Politics* 73 (4), 1038–1050.
- Singh, P., 2015. Subnationalism and Social Development: A Comparative Analysis of Indian States. *World Politics* 67 (03), 506–562.
- Singh, P., 2016. *How Solidarity Works for Welfare: Subnationalism and Social Development in India*. Cambridge University Press, New York.
- Tajima, Y., Samphantharak, K., Ostwald, K., 2018. Ethnic Segregation and Public Goods: Evidence from Indonesia. *American Political Science Review* 112 (3), 1–17.
- Trounstine, J., 2016. Segregation and inequality in public goods. *American Journal of Political Science* 60 (3), 709–725.
- Zacharias, A., Vakulabharanam, V., 2011. Caste Stratification and Wealth Inequality in India. *World Development* 39 (10), 1820–1833.

Table 1: Public Goods, Diversity, Segregation, and Hierarchy

	Fractionalization	Segregation	OTH-SCST	SCST-OTH	SCST-SCST
<i>Educational Facilities</i>					
Primary School	0.05*** (0.00)	0.02** (0.01)	-0.00 (0.00)	-0.01*** (0.00)	-0.00 (0.00)
Middle School	0.05*** (0.00)	-0.02 (0.01)	0.01 (0.00)	-0.04*** (0.00)	-0.03*** (0.00)
Secondary School	-0.03*** (0.00)	-0.02** (0.01)	0.02*** (0.00)	-0.03*** (0.00)	-0.03*** (0.00)
<i>Health Facilities</i>					
Primary Health Center	-0.02*** (0.00)	0.01* (0.00)	0.01*** (0.00)	-0.00 (0.00)	-0.01** (0.00)
Maternal & Child Welfare Center	-0.02*** (0.00)	-0.00 (0.00)	0.01*** (0.00)	-0.01*** (0.00)	-0.01*** (0.00)
Hospital	-0.01*** (0.00)	0.01 (0.00)	0.01*** (0.00)	-0.00 (0.00)	-0.00 (0.00)
Dispensary	-0.01*** (0.00)	0.00 (0.00)	0.01*** (0.00)	-0.01*** (0.00)	-0.01** (0.00)
Family Welfare Center	-0.02*** (0.00)	0.01* (0.00)	0.01*** (0.00)	-0.01*** (0.00)	-0.01*** (0.00)
<i>Water Facilities</i>					
Tapwater	0.03*** (0.00)	-0.08*** (0.01)	0.01 (0.00)	0.01* (0.00)	0.00 (0.00)
Well	0.02*** (0.00)	-0.02* (0.01)	0.04*** (0.00)	0.00 (0.00)	0.04*** (0.00)
Handpump	0.02*** (0.00)	0.05*** (0.01)	0.02*** (0.00)	0.01*** (0.00)	0.03*** (0.00)
<i>Sanitation Facilities</i>					
Drainage	0.05*** (0.00)	-0.10*** (0.01)	-0.01* (0.00)	-0.00 (0.00)	-0.00 (0.00)
Total Sanitation Campaign	0.00 (0.00)	-0.02** (0.01)	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)
Community Toilet Complex	-0.01*** (0.00)	-0.03*** (0.00)	0.00 (0.00)	-0.01*** (0.00)	-0.00 (0.00)
<i>Transport and Communication Facilities</i>					
Post Office	0.00 (0.00)	-0.05*** (0.01)	0.01** (0.00)	-0.03*** (0.00)	-0.03*** (0.00)
Bus	0.07*** (0.00)	-0.03** (0.01)	0.02*** (0.00)	0.00 (0.00)	0.02*** (0.00)
Auto-Taxi-Van	0.01** (0.00)	-0.04*** (0.01)	0.01* (0.00)	-0.01*** (0.00)	-0.00 (0.00)
<i>Road Facilities</i>					
National Highway	0.01* (0.00)	-0.00 (0.00)	0.01* (0.00)	-0.00 (0.00)	-0.00 (0.00)
State Highway	0.01*** (0.00)	-0.04*** (0.01)	0.01* (0.00)	-0.01** (0.00)	-0.00 (0.00)
Paved Road	0.07*** (0.00)	-0.02** (0.01)	0.02*** (0.00)	0.00 (0.00)	0.04*** (0.00)
All-weather Road	0.05*** (0.00)	-0.06*** (0.01)	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)
<i>Other Facilities</i>					
Bank-Credit	0.02*** (0.00)	0.03*** (0.01)	0.00 (0.00)	-0.02*** (0.00)	0.00 (0.00)
PDS	0.06*** (0.00)	-0.06*** (0.01)	0.00 (0.00)	-0.03*** (0.00)	-0.01 (0.00)
Nutrition Center	0.05*** (0.00)	0.03*** (0.01)	-0.00 (0.00)	0.00** (0.00)	0.02*** (0.00)
Electricity	0.04*** (0.00)	-0.00 (0.01)	0.01** (0.00)	0.03*** (0.00)	0.02*** (0.00)
N	595,906				

Note: Dependent Variable – incidence of respective public good (as recorded in the first column).  
+  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

*Control variables:* Village & Sub-district level area and population; share of SCs and STs, literacy rates, sex ratio, and workforce participation rate and sub-district fractionalization.

PDS refers to the Public Distribution System.

Table 2: Impact of Ethnic Hierarchy in Segregated Sub-districts

	Fractionalization	
	SCST-OTH	OTH-SCST
<i>Educational Facilities</i>		
Primary School	0.09 (0.07)	0.01 (0.02)
Middle School	0.31*** (0.09)	0.04 (0.02)
Secondary School	0.21** (0.07)	-0.01 (0.02)
<i>Health Facilities</i>		
Primary Health Center	0.08 (0.05)	0.00 (0.01)
Maternal & Child Welfare Center	0.09* (0.05)	0.02* (0.01)
Hospital	0.03 (0.04)	0.00 (0.01)
Dispensary	-0.01 (0.04)	0.00 (0.01)
Family Welfare Center	0.00 (0.04)	0.00 (0.01)
<i>Water Facilities</i>		
Tapwater	0.21** (0.08)	-0.01 (0.02)
Well	-0.04 (0.07)	0.08*** (0.02)
Handpump	-0.03 (0.06)	0.01 (0.02)
<i>Sanitation Facilities</i>		
Drainage	0.16* (0.02)	0.02 (0.02)
Total Sanitation Campaign	0.04 (0.06)	-0.01 (0.02)
Community Toilet Complex	0.05 (0.04)	-0.00 (0.01)
<i>Transport and Communication Facilities</i>		
Post Office	0.24** (0.08)	0.00 (0.02)
Bus	0.42*** (0.09)	0.02 (0.02)
Auto-Taxi-Van	0.18* (0.08)	0.03 (0.02)
<i>Road Facilities</i>		
National Highway	0.11* (0.05)	0.01 (0.01)
State Highway	0.22** (0.07)	0.04* (0.02)
Paved Road	0.27*** (0.08)	0.05* (0.02)
All-weather Road	0.00 (0.09)	0.03 (0.03)
<i>Other Facilities</i>		
Bank-Credit	0.02 (0.09)	0.01 (0.02)
PDS	0.19* (0.09)	0.02 (0.02)
Nutrition Center	0.03 (0.07)	0.01 (0.02)
Electricity	0.01 (0.06)	0.06*** (0.02)
N	13496	25780

Note: Dependent Variable – incidence of respective public good (as recorded in the first column).  
+  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

*Control variables:* Village-level share of SCs and STs, literacy rates, sex ratio, and workforce participation rate.

PDS refers to the Public Distribution System.

Table 3: Association between fractionalization and public goods

	Fractionalization
<i>Educational Facilities</i>	
Primary School	0.05*** (0.00)
Middle School	0.05*** (0.00)
Secondary School	-0.02*** (0.00)
<i>Health Facilities</i>	
Primary Health Center	-0.02*** (0.00)
Maternal & Child Welfare Center	-0.01*** (0.00)
Hospital	-0.01*** (0.00)
Dispensary	-0.01*** (0.00)
Family Welfare Center	-0.02*** (0.00)
<i>Water Facilities</i>	
Tapwater	0.03*** (0.00)
Well	0.02*** (0.00)
Handpump	0.02*** (0.00)
<i>Sanitation Facilities</i>	
Drainage	0.04*** (0.00)
Total Sanitation Campaign	-0.00 (0.00)
Community Toilet Complex	0.01*** (0.00)
<i>Transport and Communication Facilities</i>	
Post Office	0.07*** (0.00)
Bus	0.02*** (0.00)
Auto-Taxi-Van	0.01*** (0.00)
<i>Road Facilities</i>	
National Highway	0.01*** (0.00)
State Highway	0.06*** (0.00)
Paved Road	0.04*** (0.00)
All-weather Road	0.03*** (0.00)
<i>Other Facilities</i>	
Bank-Credit	0.06*** (0.00)
PDS	0.04*** (0.00)
Nutrition Center	0.04*** (0.00)
Electricity	0.04*** (0.00)
N	595,906
<i>Note: Dependent Variable - indicator variable for each facility in the row</i>	
<i>Controls: Village-level Share SC, ST, Literacy Rate, Sex Ratio, Workforce Participation Rate, Tehsil fixed effects</i>	

Table 4: Diversity, Segregation and Public Goods: Village-level Regressions

	Fractionalization	Segregation
<i>Educational Facilities</i>		
Primary School	0.06*** (0.00)	0.02** (0.01)
Middle School	0.05*** (0.00)	-0.02* (0.01)
Secondary School	-0.02*** (0.00)	-0.02** (0.01)
<i>Health Facilities</i>		
Primary Health Center	-0.02*** (0.00)	0.01** (0.00)
Maternal & Child Welfare Center	-0.01*** (0.00)	-0.00 (0.00)
Hospital	-0.00*** (0.00)	0.01* (0.00)
Dispensary	-0.01*** (0.00)	0.00 (0.00)
Family Welfare Center	-0.02*** (0.00)	0.01** (0.00)
<i>Water Facilities</i>		
Tapwater	0.03*** (0.00)	-0.08*** (0.01)
Well	0.01*** (0.00)	-0.02** (0.01)
Handpump	0.01*** (0.00)	0.05*** (0.01)
<i>Sanitation Facilities</i>		
Drainage	0.05*** (0.00)	-0.10*** (0.01)
Total Sanitation Campaign	0.00 (0.00)	-0.02** (0.01)
Community Toilet Complex	-0.01*** (0.00)	-0.03*** (0.00)
<i>Transport and Communication Facilities</i>		
Post Office	0.01*** (0.00)	-0.05*** (0.01)
Bus	0.07*** (0.00)	-0.03*** (0.01)
Auto-Taxi-Van	0.01*** (0.00)	-0.04*** (0.01)
<i>Road Facilities</i>		
National Highway	0.01** (0.00)	0.00 (0.00)
State Highway	0.01*** (0.00)	-0.04*** (0.01)
Paved Road	0.06*** (0.00)	-0.03** (0.01)
All-weather Road	0.05*** (0.00)	-0.06*** (0.01)
<i>Other Facilities</i>		
Bank-Credit	0.02*** (0.00)	0.03*** (0.01)
PDS	0.06*** (0.00)	-0.06*** (0.01)
Nutrition Center	0.05*** (0.00)	0.03*** (0.01)
Electricity	0.04*** (0.00)	0.00 (0.01)
No. of villages	595906	

Note: Dependent Variable – indicator variable for each facility in the row.

+  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

*Control variables:* Village & Sub-district level area and population; share of SCs and STs, literacy rates, sex ratio, and workforce participation rate and sub-district fractionalization.

PDS refers to the Public Distribution System.

Table 5: Nature of association between diversity and public goods by segregation quartiles

	Quartile I	Quartile II	Quartile III	Quartile IV
<i>Educational Facilities</i>				
Primary School	+	+	+	+
Middle School	+	+	+	+
Secondary School	-	-	+	+
Senior Secondary School	-	-	-	+
<i>Health Facilities</i>				
Primary Health Center	-	-	0	+
Maternal & Child Welfare Center	-	-	0	+
Hospital	-	-	0	+
Dispensary	-	-	0	+
Family Welfare Center	-	-	0	0
<i>Water Facilities</i>				
Tapwater	+	0	0	+
Well	+	0	+	+
Handpump	0	0	+	+
<i>Sanitation Facilities</i>				
Drainage	+	0	+	+
Total Sanitation Campaign	0	0	-	0
Community Toilet Complex	-	-	-	+
<i>Transport and Communication Facilities</i>				
Post Office	-	0	+	+
Bus	+	+	+	+
Auto-Taxi-Van	-	0	+	+
<i>Road Facilities</i>				
National Highway	-	+	+	+
State Highway	-	0	+	+
Paved Road	+	+	+	+
All-weather Road	+	+	+	+
<i>Other Facilities</i>				
Bank-Credit	0	0	0	0
PDS	+	+	+	+
Nutrition Center	+	+	+	0
Electricity	+	+	+	+
No. of villages	595906			

Note: Dependent Variable – indicator variable for each facility in the row.

+ refers to positive and significant coefficient, – refers to negative and significant coefficient, and 0 refers to insignificant coefficient on fractionalization.

*Control Variables:* Village level area, total population, share of SCs and STs, literacy rates, sex ratio, and workforce participation rate. We have also controlled for sub-district level fixed effects.

PDS refers to the Public Distribution System.

Detailed results are available upon request.

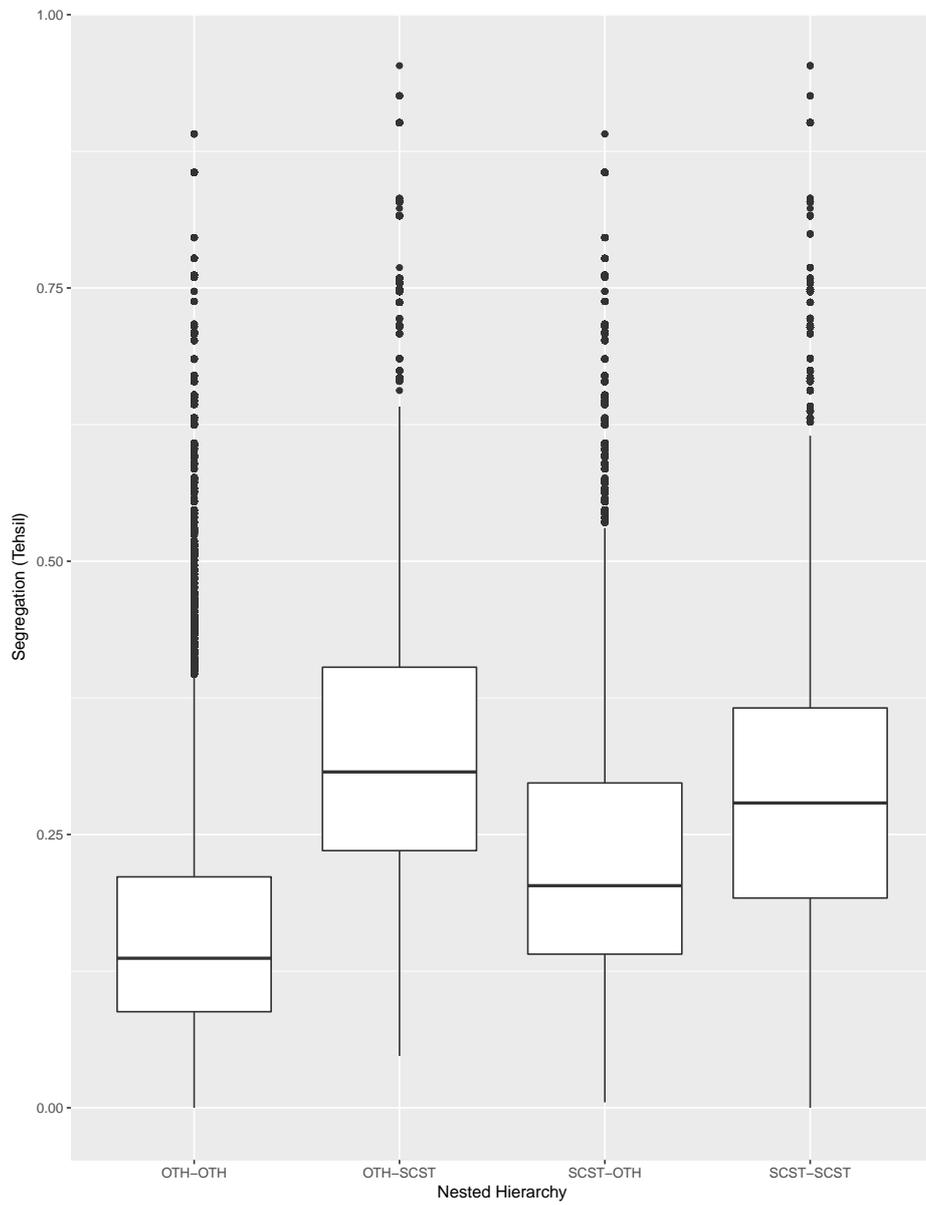


Figure 4: Segregation and Hierarchy

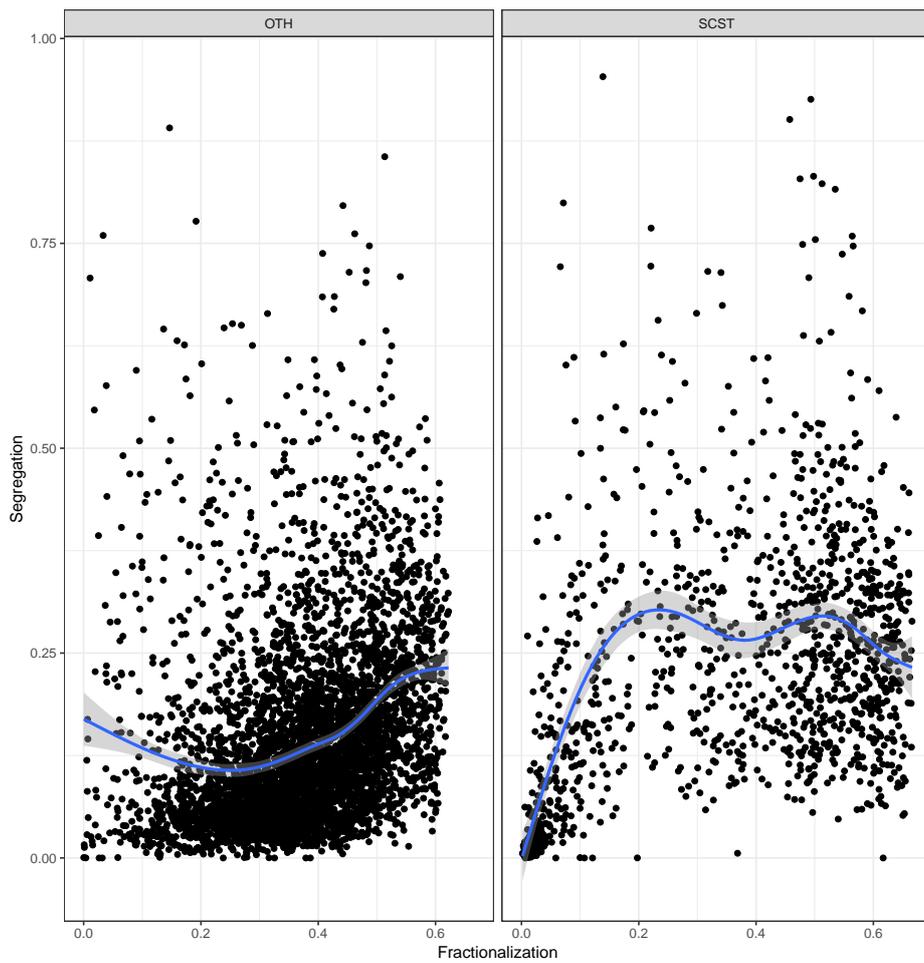


Figure 5: *Fractionalization, Segregation, and Hierarchy*