WP 2018-07

May 2018

Working Paper

Dyson School of Applied Economics and Management
Cornell University, Ithaca, New York 14853-7801 USA

Spatial Segregation, Multi-scale Diversity, and Public Goods

Naveen Bharathi, Deepak Malghan, Sumit Mishra and Andaleeb Rahman

.

It is the policy of Cornell University actively to support equality of educational and employment opportunity. No person shall be denied admission to any educational program or activity or be denied employment on the basis of any legally prohibited discrimination involving, but not limited to, such factors as race, color, creed, religion, national or ethnic origin, sex, age or handicap. The University is committed to the maintenance of affirmative action programs which will assure the continuation of such equality of opportunity.

Spatial Segregation, Multi-scale Diversity, and Public Goods

Naveen Bharathi[a], Deepak Malghan[a], Sumit Mishra[b], Andaleeb Rahman[*,c]

[a] *Indian Institute of Management Bangalore, Bengaluru 560 076, INDIA*
[b] *Institute for Financial Management and Research, Chennai, 600 034, INDIA*
[c] *Cornell University, Ithaca NY 14853, USA*

Abstract

We develop a general multi-scale diversity framework to account for spatial segregation of ethnic groups in politically nested geographic aggregations. Our framework explains why the celebrated "diversity-debit hypothesis" in political economy of public goods is sensitive to spatial unit of analysis, and how not accounting for segregation biases empirical diversity-development models. We test our framework using census data from Indian villages ($n \approx 600,000$) and sub-districts containing these villages ($n \approx 6,000$), for twenty-five different public goods.

Key words: Ethnic Diversity, Segregation Index, Nested Spatial Scales, Public Goods Catchment Area

[*] Corresponding author. Telephone: +1 607 280 5881.
Email addresses: naveen.bharathi@iimb.ac.in (Naveen Bharathi), dmalghan@iimb.ac.in (Deepak Malghan), sumit.mishra@ifmr.ac.in (Sumit Mishra), ar687@cornell.edu (Andaleeb Rahman)

May 28, 2018

1. Introduction

Ethnic, linguistic, racial, or religious diversity is negatively associated with provision of public goods and other development outcomes (Alesina and La Ferrara, 2005). This negative relationship is driven by varying preferences and coordination failures, if not outright strife between ethnic groups (Habyarimana et al., 2007). While recent empirical evidence suggests a breakdown of the "diversity debit hypothesis" at subnational scales, little is understood about why the diversity-development relationship might be sensitive to spatial scale (Gerring et al., 2015; Gisselquist et al., 2016).

In this paper, we show that the diversity-development relationship is not only driven by intra-unit heterogeneity but also by spatial segregation of politically salient ethnic groups. In particular, we argue that when a geographic unit is hierarchically nested within multiple administrative and political aggregations involved in public goods provisioning, spatial distribution of ethnic groups across these geographic aggregations is an important determinant of public goods outcomes. For example, village-level public goods are not only impacted by local ethnic diversity but also by patterns of diversity within the sub-district that contains the village, the district that contains the sub-district, and finally the sub-national province containing the district. We use the latest available census data from rural India (2011) to empirically demonstrate the salience of this spatial segregation channel. Using data for twenty-five different public goods from nearly six hundred thousand villages across India, we empirically establish that a multi-scale diversity-context rather than simple intra-unit diversity determines public goods provisioning.

2. Multi-scale Diversity Context, and Segregation

Figure 1 shows why a multi-scale diversity context must account for segregation as much as intra-unit diversity. The figure represents 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, ethnic segregation is different 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, primary school will have a smaller catchment than high schools. We represent the multi-scale diversity-context relevant for provisioning of P at spatial-scale z that is nested in Z as:

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

Extant literature has neglected both inter-unit segregation, and catchment area as factors relevant

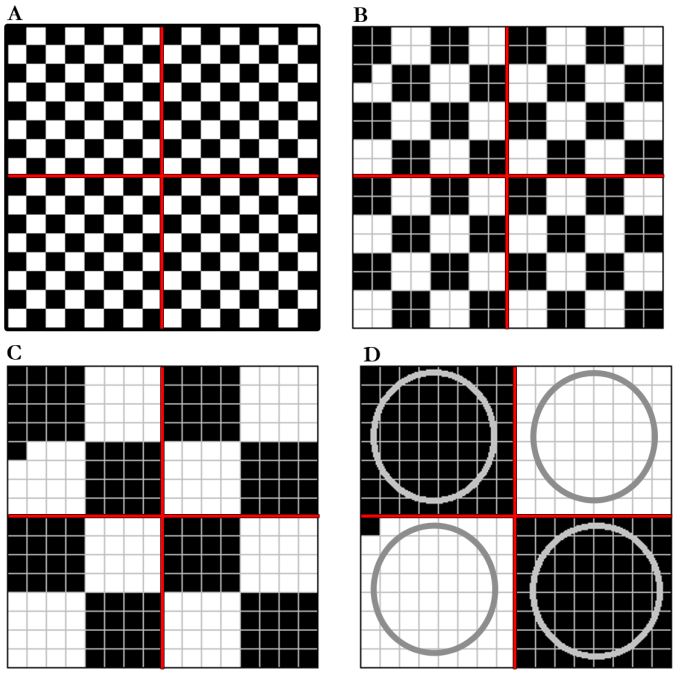

Figure 1: *Segregation, Catchment Area, and Multi-scale Diversity Context*

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 happens at higher echelons of the state, lower-level state actors 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, and geographic segregation abets both these channels. When ethnic diversity is measured at large geographic aggregates – as is the norm in extant literature – politically salient micro-ecologies of local segregation 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 is particularly important in nested political contexts that we have described here – homogeneous neighbourhoods (villages) can exist within diverse higher-order geographic aggregations (district or sub-districts) 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. 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 leads to higher political polarization inhibiting cooperation needed for effective public good demand (Trounstine, 2016). It is easier for homogeneous geographic units to organize and politically articulate their 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).

Finally, accounting for spatial segregation is also important for statistical inference. In a nested geography, statistical interpretations suffer from the modifiable areal unit problem, or MAUP (Openshaw, 1984). MAUP is a spatial version of a more general statistical inference problem – the ecological fallacy.

3. Evidence from Indian Census Data

We illustrate how a multi-scale diversity context that includes segregation information impacts public goods provisioning using data from census of all villages in India (2011). India has a three-tier government – federal/union level, state-level, and local-level. 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 form clusters of villages.[1] Ethnic favouritism has often 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).

The most important social cleavage in rural India is that of caste. 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 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).

In their seminal work on public goods in rural India, 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 parliamentary constituency level. Parliamentary constituencies are large aggregates that can contain over a thousand villages and pose problems of statistical inference related to MAUP in addition to neglecting the nested diversity context. In order to overcome these theoretical and empirical limitations, we use the most elementary administrative unit – the village – as our spatial unit of analysis.

[1] The 595,906 villages in our dataset are clustered into 238,617 *panchayats* for an average of ≈ 2.5 villages per *panchayat*.

3.1. Data and Diversity Metrics

We use the 2011 national census data which contains caste information (aggregated into three politically salient categories) at the village level ($n = 595,906$) that we aggregate into sub-districts ($n = 5878$) to construct a segregation index. The national census (village directory data) also contains incidence information for several public goods that we use as our dependent variables. This data covers a diverse array of public goods including education, health, infrastructure, and transport (Table 1 contains a full list of our dependent variables).

We construct the fractionalization metric (ELF) for all villages – the workhorse metric used in diversity-development literature – that represents the probability that two randomly chosen individuals belong to distinct social groups. Using the notation introduced in Section 2, for any village $i \in z$, fractionalization index is simply:

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

where $k \in \{SC, ST, OTH\}$ represents census-designated subgroup in village i, and π_{ik} is the population share of subgroup k in village i.[2] 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)

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

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 aggregate fractionalization index computed at the sub-district level. The spatial distribution of sub-district segregation is shown in Figure 2. The figure shows quartiles with an cumulative density function as an inset. The distribution has a sparse and long right tail so that for over 95% of the sub-districts, $SEG_j \leq 0.5$.

4. Results

In order to investigate the effect of spatial segregation on public goods provisioning, we estimate a village-level linear probability model (LPM) of the following form:

$$Y_{ip} = \alpha_i + \beta_i \times FRA_i + \gamma_j \times SEG_j + \vec{\theta_i} \cdot \vec{V_i} + \vec{\delta_j} \cdot \vec{T_j} + \epsilon_i \tag{4}$$

where Y_{ip} is the incidence of public good p in village i (contained within sub-district j). Besides district level fixed effects, we include village-level ($\vec{V_i}$), and sub-district level ($\vec{T_j}$) controls vectors. Coefficients on FRA_i and SEG_j from this regression is presented in Table 1. Contrary to many of the existing findings, we find a clear evidence for diversity-dividend at village-level for most public goods except for the provision of secondary schools and health facilities. However, sub-district level segregation lowers the probability of a village having access to middle and secondary schools, water facilities (tap and well), sanitation facilities, all transport and communication facilities, road

[2] Census records social groups as Scheduled Castes (SC); Scheduled Tribes (ST); and Others (OTH).

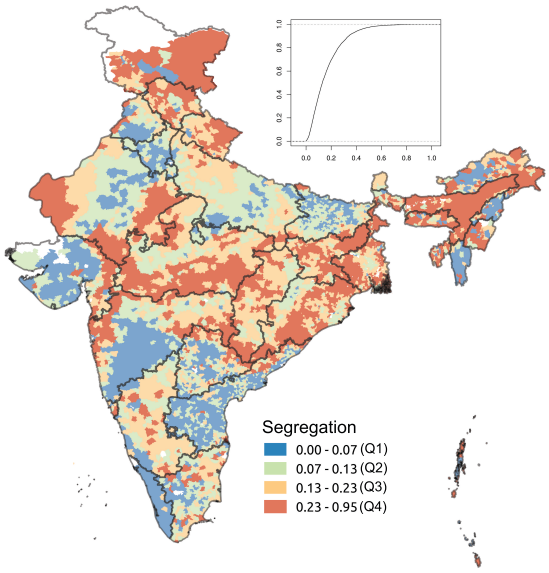

Figure 2: *Sub-district Segregation Map. Inset shows cumulative density function.* $n = 5878$.

facilities (except for national highway), public distribution (PDS). On the other hand, sub-district level segregation is positively associated with village having access to nutrition center, bank or credit facility, health facilities, and primary school.

4.1. Segregation Quartile Analysis

To further explicate the role of spatial segregation, we divided our village data into four sub-samples corresponding to the sub-district segregation quartiles (Figure 2), and estimated the same models reported in Table 1 for each of the four sub-samples, with a sub-district fixed effect included. Table 2 reports the sign and significance of the coefficients on fractionalization (FRA_i) for each segregation quartile. Villages in sub-districts with low 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.

5. 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 empirical analysis of the relationship between diversity and development.

References

Alesina, A., La Ferrara, E., 2005. Ethnic Diversity and Economic Performance. Journal of Economic Literature 43 (3), 762–800.

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.

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 Dev. 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.

Munshi, K., Rosenzweig, M., 2016. Insiders and Outsiders: Local Ethnic Politics and Public Goods Provision, Working Paper.

Openshaw, S., 1984. Ecological fallacies and the analysis of areal census data. Environment and planning A 16 (1), 17–31.

Tajima, Y., Samphantharak, K., Ostwald, K., 2018. Ethnic Segregation and Public Goods: Evidence from Indonesia. American Political Science Review (1999), 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: *Diversity, Segregation and Public Goods: Village-level Regressions*

	Fractionalization	**Segregation**
Educational Facilities		
Primary School	0.06***	0.02**
	(0.00)	(0.01)
Middle School	0.05***	-0.02*
	(0.00)	(0.01)
Secondary School	-0.02***	-0.02**
	(0.00)	(0.01)
Health Facilities		
Primary Health Center	-0.02***	0.01**
	(0.00)	(0.00)
Maternal & Child Welfare Center	-0.01***	-0.00
	(0.00)	(0.00)
Hospital	-0.00***	0.01*
	(0.00)	(0.00)
Dispensary	-0.01***	0.00
	(0.00)	(0.00)
Family Welfare Center	-0.02***	0.01**
	(0.00)	(0.00)
Water Facilities		
Tapwater	0.03***	-0.08***
	(0.00)	(0.01)
Well	0.01***	-0.02**
	(0.00)	(0.01)
Handpump	0.01***	0.05***
	(0.00)	(0.01)
Sanitation Facilities		
Drainage	0.05***	-0.10***
	(0.00)	(0.01)
Total Sanitation Campaign	0.00	-0.02**
	(0.00)	(0.01)
Community Toilet Complex	-0.01***	-0.03***
	(0.00)	(0.00)
Transport and Communication Facilities		
Post Office	0.01***	-0.05***
	(0.00)	(0.01)
Bus	0.07***	-0.03***
	(0.00)	(0.01)
Auto-Taxi-Van	0.01***	-0.04***
	(0.00)	(0.01)
Road Facilities		
National Highway	0.01**	0.00
	(0.00)	(0.00)
State Highway	0.01***	-0.04***
	(0.00)	(0.01)
Paved Road	0.06***	-0.03**
	(0.00)	(0.01)
All-weather Road	0.05***	-0.06***
	(0.00)	(0.01)
Other Facilities		
Bank-Credit	0.02***	0.03***
	(0.00)	(0.01)
PDS	0.06***	-0.06***
	(0.00)	(0.01)
Nutrition Center	0.05***	0.03***
	(0.00)	(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.
Detailed results are available upon request.

Table 2: Nature of association between diversity and public goods by segregation quartiles

	Quartile I	Quartile II	Quartile III	Quartile IV
Educational Facilities				
Primary School	+	+	+	+
Middle School	+	+	+	+
Secondary School	-	-	+	+
Senior Secondary School	-	-	-	+
Health Facilities				
Primary Health Center	-	-	0	+
Maternal & Child Welfare Center	-	-	0	+
Hospital	-	-	0	+
Dispensary	-	-	0	+
Family Welfare Center	-	-	0	0
Water Facilities				
Tapwater	+	0	0	+
Well	+	0	+	+
Handpump	0	0	+	+
Sanitation Facilities				
Drainage	+	0	+	+
Total Sanitation Campaign	0	0	-	0
Community Toilet Complex	-	-	-	+
Transport and Communication Facilities				
Post Office	-	0	+	+
Bus	+	+	+	+
Auto-Taxi-Van	-	0	+	+
Road Facilities				
National Highway	-	+	+	+
State Highway	-	0	+	+
Paved Road	+	+	+	+
All-weather Road	+	+	+	+
Other Facilities				
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.

OTHER A.E.M. WORKING PAPERS

WP No	Title	Fee (if applicable)	Author(s)
2018-07	Spatial Segregation, Multi-scale Diversity, and Public Goods		Bharathi, N., Malghan, D., Mishra, S. and A. Rahman
2018-06	The Compatibility of Farmland Sales and Opinion Survey Data		Bigelow, D.P., Ifft, J. and T. Kuethe
2018-05	Exploring the Feasibility of Rural Broadband Cooperatives: The NEW New Deal?		Schmit, T. M. and R. M. Severson
2018-04	Employer Power, Labor Saving Technical Change, and Inequality		Chau, H.N. and R. Kanbur
2018-03	On the Volume of Redistribution: Across Income Levels and Across Groups		Kanbur, R.
2018-02	Gunnar Myrdal and *Asian Drama* in Context		Kanbur, R.
2018-01	Political Activism as a Determinant of Clientelistic Transfers: Evidence from and Indian Public Works Program		Chau, N., Liu, Y. & V. Soundararajan
2017-14	Too Much to Eat It All: How Package Size Impacts Food Waste		Petit, O., Lunardo, R. and B. Rickard
2017-13	Why Secondary Towns Can Be Important For Poverty Reduction - A Migrant's Perspective		Ingalaere, B., Christiaensen, L., De Weerdt, J. and R. Kanbur
2017-12	Citizenship, Migration and Opportunity		Kanbur, R.
2017-11	The Digital Revolution and Targeting Public Expenditure for Poverty Reduction		Kanbur, R.
2017-10	What is the World Bank Good For? Global Public Goods and Global Institutions		Kanbur, R.
2017-09	Sub-Saharan Africa's Manufacturing Sector: Building Complexity		Bhorat, H., Kanbur, R., Rooney, C. and Steenkamp, F.
2017-08	Farmer Productivity By Age Over Eight U.S. Census Years		Tauer, L.W.
2017-07	Inequality Indices as Tests of Fairness		Kanbur, R. and Snell, A.

Paper copies are being replaced by electronic Portable Document Files (PDFs). To request PDFs of AEM publications, write to (be sure to include your e-mail address): Publications, Department of Applied Economics and Management, Warren Hall, Cornell University, Ithaca, NY 14853-7801. If a fee is indicated, please include a check or money order made payable to Cornell University for the amount of your purchase. Visit our Web site (*http://dyson.cornell.edu/research/wp.php)* for a more complete list of recent bulletins.