Working Paper

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

Village in the City: Residential Segregation in Urbanizing India

Naveen Bharathi, Deepak Malghan, 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.
Residential Segregation in Urbanizing India

Naveen Bharathi (naveen.bharathi@iimb.ac.in)
Deepak Malghan (dmalghan@iimb.ac.in)
Andaleeb Rahman (ar687@cornell.edu)

Abstract

One of the normative promises of Indian urbanization is the potential breaking down of the rigidities that characterize traditional caste hierarchies in an agrarian regime. In particular, urbanization holds the promise of breaking down spatial barriers between traditional caste groups. Using a unique census-scale dataset from urban Karnataka containing detailed caste and religion data, we present a contemporary snapshot of the relationship between urbanization and patterns of residential segregation.

Our analysis shows that urban wards (the extant elementary spatial unit used in the literature) are heterogeneous and segregation within the wards at census-block scales account for a significant part of the overall patterns of city scale segregation. In particular, we show how intra-ward segregation is a central driver of ghettoization of the most spatially marginalized groups in urban India – Muslims and Dalits. We provide the first census-scale evidence in independent India that corroborates anecdotal accounts of urban ghettoization.

The cross-section snapshot presented in this paper suggests that degree of residential segregation is uncorrelated with levels of urbanization. We report high levels of segregation across a diverse set of urban centers that include semi-urban settlements to arguably India’s most globalized metropolis of over ten million residents, Bengaluru.

NOTE: This paper uses material first reported in our 2018 working paper titled “Isolated by Caste: Neighborhood-Scale Residential Segregation in Indian Metros”.
Introduction

Caste in India is one of the most complex and rigid social structures whose basic characteristics include complex network of hierarchies, segmentation, and segregation (Ghurye, 1969). An individual’s caste marker is closely correlated with both class and power (Beteille, 1967). Spatial segregation and discrimination have been one of the basic aspects of the caste system (A. R. Desai, 1994). Many founding fathers of the Indian Constitution, advocated greater urbanization for Dalits (previously known as ‘untouchables’, and administratively referred to as the ‘Scheduled Castes’) as a solution to escape the most egregious consequences of caste. The structure of caste in the process of urbanization is expected to fuse with class blurring the inherent stratification of caste. It has been argued that social status in urban Indian is characterized by markers of class (such as income, wealth, and education) rather than caste (Beteille 1997). Scholars have previously observed that caste rules do no control the organization of spatial environment of a city as they do in a village (Swallow, 1982). Contemporary Dalit intellectuals have suggested that in an urbanizing India, ‘caste is losing, and will continue to lose, its strength’ (C. B. Prasad, 2010). Historically, communities have migrated to cities to ‘escape’ caste hegemony in the villages – Mahars (an ‘untouchable’ caste) migrated to cities like Bombay and Nagpur in large numbers at the beginning of the 20th century (Rao, 2009). Also, the population of Dalits in Urban India saw an increase of 40 percent in the decade 2001-2011 (National Census, 2011). A study of Bangalore slums by Krishna, Sriram, & Prakash (2014), reports a disproportionately large share of people belonging to Scheduled Castes (around 72 percent as compared to
11.41 percent for overall) in newer slums which indicates large migration of Dalits from villages to cities.

In contemporary India, caste continues to have a major impact on various socio-economic outcomes like education, health, labor markets and electoral politics (Borooah, 2010; Kothari, 1995; Nambissan, 2009; S. Thorat & Neuman, 2012). Rights, access, citizenship and privileges of an individual are often tied to his or her membership of a particular caste (Deshpande, 2000, 2001; Thorat, Banerjee, Mishra, & Rizvi, 2015; Thorat & Neuman, 2012b). Communities belonging to lower castes are socially excluded, marginalized and often denied basic human rights, freedom and dignity (Kothari, 1994). The structure of caste in the Indian subcontinent is not specific to Hinduism but permeates other religions too. Hence, caste is the ultimate social and individual attribute in India.

The republican constitution of India, which came into effect in 1950, abolished the practice of untouchability (Article 17). However, there are no provisions in the Constitution that abolish the caste system itself, or its spatial manifestations. Thorat and Joshi (2015) use IHDS data to report that in 2011-12, twenty-percent of all urban households and 30 percent of all rural households practiced some, or the other form of untouchability. Five percent of Dalit households also practiced untouchability – showing the highly hierarchical nature of caste which thrives on discrimination and division. This phenomenon cut across religions – Jains (35%), Hindus (30%), Islam (18%), Sikh (23%), and Christian (5%) – which is testament to its resilient nature -- as in religions such as Islam, Sikhism and Christianity there is no concept of caste in these religions. In spite of this stark reality the debates on caste and untouchability in urban spaces are almost absent in public debates and media (Guru & Sarukkai, 2014). Caste, in today’s scenario,
might not be that obvious but it ‘lurks somewhere beneath the surface’ and ‘caste will always be there’.

Dirks (2003) while documenting transformation colonialism brought about on the structure of caste, argues that in contemporary society, ‘sociological assurance that caste would disappear except as a form of domestic ritual or familial identity when it entered the city and new domains of industrial capital turned out to be a bourgeois dream disrupted both by steady reports of escalating caste violence in the countryside and then the turmoil over reservations in the principal cities of the nation’.

The dominant strands in Indian urbanism have however not studied caste as a significant factor influencing the politics of space making (Nair, 2013). Instead, urban scholars of Indian cities have focused on class markers despite evidence that caste implicated in access to economic resources and opportunities even in modern sectors of India (Thorat & Neuman, 2012a). Even if caste is not an essential feature of Indian society (Dumont, 1966), and is not the fountainhead of other identities, critical scholars have argued that the most pernicious features of casteism including segregation can be overcome only by ‘caste action’ and not ‘class action’ (Omvedt, 1978).

Urbanization and Caste

One of the primary aspects of caste system is residential segregation (Ghurye, 1969). In rural India, caste groups at the bottom of caste hierarchy typically reside on the outskirts of the village. The central parts of the village continue to be segregated along caste lines corresponding to occupational and ritual hierarchical status. Such segregated configurations determine caste groups’ access to public goods such as village well, grazing fields, etc. Caste hierarchies and spatial segregation are mutually reinforcing so

---

1 Shashi Tharoor in ‘Why Caste Won’t Disappear From India’ in Huffington Post, Dated: 09/12/2014
that residential segregation is a both a product of hierarchical relationship between caste groups, and a key contributor to strengthening and persistence of such hierarchies. If the social distance and hierarchies of various castes are reflected in the spatial segregation of residential localities in a settlement (Mukherjee, 1968), social hierarchies are in turn reinforced by spatial isolation and separation. Segregation reduces the likelihood of any social interaction across social groups. Muslims in Indian cities are a classic example of such isolation (Gayer & Jaffrelot, 2012). Instead, contacts tend to be formalized, confined principally to the market place or work place (Hazlehurst, 1970). Such contacts are usually marginal in value and are not socially rich. People who work together or who have contacts of a strictly economic character may live in entirely different social and ecological worlds (Gist & Fava, 1970).

The received wisdom, or more accurately hope, is that the sense of anonymity provided by urban areas should make caste based segregation rarer as urbanization progresses. Historical evidence however suggests that Indian cities have also been segregated along caste lines. According to Karim (1956) the city population in pre-British India was largely segregated geographically by religion, caste and sub-caste, and by occupational and regional groups forming social islands. These social units developed such exclusiveness that the groups constituted cities within cities. Gist (1957) in his study of Bangalore documents caste and religion based residential segregation. Studies of *jataw thoks* and *mohollas* of Agra by Lynch(1967) and of Lucknow *rickshawallas* by Gould(1965) show caste as a predominant factor in the organization of urban neighborhoods. Similarly, Hazlehurst (1970) in his study of Puranapur in Haryana finds that while public spaces like markets showed a mixture of castes, residential neighborhoods were highly segregated along caste lines.
Mainstream urban theorists broadly concur that industrialization and ensuing urbanization are expected to provide equal opportunities to individuals of different backgrounds to progress and thrive. Urbanization is considered an indicator of economic growth and material development. Cities are the place of the ‘individual’. A well-known statement of this abiding hypothesis of social change is German sociologist Ferdinand Tönnies’ (1887) distinction between *Gesellschaft* and *Gemeinschaft*. If *Gesellschaft* represents the modern industrial urban society where traditional bonds of family, kinship, caste and religion are weakened by individualism, *Gemeinschaft* refers to traditional rural societies with strong communal and familial bonds. In the *Gesellschaft*, human relations are guided by rational and utilitarian motives rather than traditional values. Another German rationalization sociologist, Simmel (1950) argues that market economy and multiple rigid bureaucratic organizations ‘rationalize’ and ‘depersonalize’ the urban community. In an urban industrial society, social relations are mediated by impersonal money economy with its calculations of profit and loss.

Chicago school urbanists who were influenced by Simmel consider a city as a key determinant for ‘social action’ and urbanism as a way of life characterized by secularization, secondary group relationships, and poorly defined social norms (Wirth, 1938). Wirth defined a city as ‘relatively large, dense and permanent settlement of socially heterogeneous individuals.’ In a city, inter-personal relationships are relationships of utility creating a scope for the disintegration of primordial caste and communal ties. Social interaction with multiple cultures and personalities loosens the grip of caste ties; class structure becomes complex. Loosening of caste and class ties increases social mobility, and social mobility increases physical mobility leading to diversity. In an industrial city, the residence patterns don’t reflect the differences or
hierarchies as compared to pre-industrial cities. Wirth also stressed that the large population size, density, and heterogeneity were important factors which produced urbanism. The bigger the city is, the more diverse it is and density makes people of different identities live together, leading to greater tolerance levels. Chicago school urbanists were in a way pioneers in studying spatial patterns to understand social phenomena.

Few scholars have challenged this near-utopian view of an urban setting. Lewis (1963) criticizes these theories which see rural and urban in dichotomous terms and for focusing on the city as a source of change. William Kolb (1954) in ‘The Social Structure and Functions of Cities’ opines that urbanization necessarily need not create the primacy of secondary relations and isolation. Morris (1968) feels that the romantic idea of city which Wirth and others created neither corresponds to historical cities nor modern urban centers. Another study of a Chicago neighborhood by Suttles (1968) shows troubled relations between various ethnic groups in the neighborhood and a strong community bond disproving the ‘atomism’ aspect of urbanization of various theories. Venkatesh (2008) in his study of Robert Taylor Homes, a ghettoized Black neighborhood in Chicago, finds a strong overlap of poverty and crime aided by strong community bonds facilitated by gangs. Gans (1962) in his study of Italians in Boston has documented strong communal bonds which were demonstrated when the settlement in which they were living was demolished and many such migrant communities have strong connections. Similarly, Whyte (1943) stresses the endurance of community ties in an American slum. Mehta (1969) in his study of Pune demonstrates that urbanization, industrialization or modernization will have no effect on residential segregation based on caste, religion or ethnicity. In India, most of the communal violence happens in urban
centers. Hansen (2001) in his study on the rise of Shiv Sena in Mumbai documents how the bonds of caste and religion got transplanted from a village and hardened in an urban setting.

In the early years of the Indian republic, Nehruvian industrial socialism was seen as one of the pathways out of the limitations of narrow parochial identities such as caste, language and religion. The expectation was that caste would stop being a hermetically sealed institution in its urban avatar and instead merge into class. But does industrialization, economic growth and urbanization dilute the strong communal and kinship bonds of caste? While robust and systematic empirical evidence is scant, anecdotal accounts suggest that the Indian caste system, rather than collapsing to class has instead adapted even while preserving its essential features. Economic liberalization in the 1990s prompted the Indian cities to remodel themselves along post-industrial globalized metropolises. Economic liberalization changed how one looked at Indian cities and at least, at the superficial level it seemed that there was a change in class and caste relations. Caste was irrelevant in a globalized setting and influence of traditional caste is being eroded by ‘forces of urbanism, secularization, and consumerism’ (D’Mello & Sahay, 2008). In a city which is dominated by private sector – where individual agency, knowledge and competence create meritocratic hierarchies where caste based ties become irrelevant. But in personal spaces caste can still have a major influence (D’Mello & Sahay, 2008). Barbara Harriss-White’s (2003) work on rural Tamil Nadu demonstrates that India’s foray into globalization has not reduced caste based inequalities.

Desai and Dubey (2011) in their nationwide study on caste based inequalities show that early urbanization benefitted privileged upper caste groups. In metropolitan
areas this discrimination is somewhat moderate but caste disparities do not disappear with urbanization. The same study demonstrates that caste inequalities are low in developed villages and metropolitan cities compared to smaller cities. Miklian and Sahoo (2016) in their study of three Indian cities note that rather than being “melting pots” and places for upward social mobility, Indian cities stubbornly mirror India’s rural social and economic realities’ and rural social structures are often repeated in urban settings. They also found out that Muslims and Dalits found it hard to escape the inherent discrimination in a metropolis and thereby pushing them to live in fringes and in segregated neighborhoods. Ahuja and Osterman (2015) while studying Indian marriage market observe that the grasp of caste in urban areas is low as compared to rural areas as urban areas provide ‘relative anonymity’ from the practices of purity and pollution.

**Residential Segregation and Caste**

Residential segregation refers to spatial separation of different groups in a given geographical area. People can be residentially segregated along various dimensions – class, race, language, religion etc. Residential segregation studied together with different levels of urbanization can indicate whether urbanization and rural to urban migration is a solution to escape caste based discrimination. The patterns and characteristics of residential segregation in US and European settlements have been extensively studied – earliest studies on the subject date back to Chicago School (e.g., (Burgess, 1928; Park, Burgess, McKenzie, & Wirth, 1925 etc.). There has been extensive study of segregation of African Americans and other ethnic groups in US cities since then. Massey and
in a major departure from the segregation studies before them, demonstrate the relationship between segregation and the creation of ‘underclass’.

Segregation brings down neighborhood level diversity. There have been two divergent arguments on whether neighborhood diversity is good for the wellbeing of the community and the individual. Putnam (2007) in ‘Bowling Alone’ felt that heterogeneity has serious negative impact on society and deteriorates community life. Further, he feels that diversity of a community and solidarity and trust between groups within that community are inversely related. Shaw & McKay (1942) state that ethnic heterogeneity undermines the ability of a community to control its members and thereby facilitating criminal behavior. Same theories of social capital are often extended to economic development – indicating that diversity might hinder economic growth.

The second set of studies refute this argument. Communities with more diversity have fewer crime rates compared to homogeneous communities (Graif & Sampson, 2009; Letki, 2008; Portes & Vickstrom, 2015). Any segregation, as research on race in US cities show, is detrimental to economic growth, societal equity, and economic mobility leading to alienation of communities (Cutler & Glaeser, 1997). A study by Chetty, Hendren, Kline, & Saez (2014) has shown that the neighborhood where one grows up has a major impact on his/her lifetime earnings and success later in life. The research found that low residential segregation results in upward social and economic mobility. Residential segregation aggravates the existing socio-economic inequality. Neighborhoods shape the lives of the children and the youth. Children growing in highly segregated poor neighborhoods are more susceptible to failure and emotionally vulnerable (Harding, 2003). Neighborhood diversity can have a positive impact on these disadvantaged groups by exposing them to mainstream role models and successful individuals (Wilson,
1987). Ludwig et al. (2013) prove that moving to a better neighborhood had a positive impact on the physical and mental health of the disadvantageous households. Putnam (2007) in fact counters his own theory by arguing that the negative effects of diversity are just temporary and the positive aspects of diversity will takeover in the long run.

Segregation also results in ghettoization of minority and poor groups and this aspect of stratification spills over to next generations (Morgan, 1984). In times of communal violence, it becomes easy to target individuals of a particular group or community. Los Angeles riots of 1992, for example, was due to highly segregated residential neighborhoods with ‘unequal social and political endowments and economic niches’ (Morrison, Lowry, & Rand Corporation, 1993). Segregation of residential areas on caste/race lines has resulted in concentration of poverty (for example slums). Residential segregation keeps intact the existing social and economic divisions and over time, can undermine prosperity (Carr & Kutty, 2008). Black families who moved to predominantly white neighborhoods achieved significant amount socio-educational gains (Rubinowitz & Rosenbaum, 2000). Ihlanfeldt & Scafidi (2002) conclude that people living in heterogeneous neighborhoods are less discriminative towards people belonging to other races and ethnic groups.

In Indian cities, often individuals are denied access to housing of their choice based on the caste that they belong (Thorat et al., 2015). Thomas Schelling(1971) in his seminal paper showed that this simple act of denial or a small preference for one’s neighbors to be of the same caste/race could lead to total segregation. Such preferences directly or indirectly will give rise to highly segregated neighborhoods and localities based on caste and often the distribution of public services and goods are decided based on the type of neighborhood (Miklian & Sahoo, 2016).
Spatial segregation in India has received scant academic attention. In particular, formal quantitative characterization of patterns of urban segregation have been lacking. Given data limitations that we describe here, ethnographic accounts have dominated the literature (for example, Gist, 1957; Hazlehurst, 1970; Lynch, 1967). Census of India reports caste information in three broad aggregate categories – SC, ST, and Others, and these broad aggregates are reported at the ward level. Thus the unit of analysis is limited in both spatial resolution as well as ethnic resolution (Dupont, 2004; Sidhwani, 2015; Vithayathil & Singh, 2012). As these papers themselves state, there are serious limitations of using ward level census data to study caste segregation in a city. The average population in an urban ward can vary from 1500 to 6000 for small towns and municipalities. In larger metropolitan cities, ward size may vary from 30000 to 200,000 (R. N. Prasad, 2006). Hence for studying neighborhood level segregation, the ward is scarcely the most useful level of analysis. A census enumeration block (or sub-block) has around 100-125 households with a population of 650-700. The geographical area of a block roughly constitutes a neighborhood – which is the optimal scale for study of residential segregation. Anecdotal evidence as well as ethnographic accounts suggest that intra-ward segregation is pervasive – especially in larger urban agglomerations such as Bengaluru that we study in detail here.

Census also does not collect detailed caste data. Even if a ward is diverse in terms of caste composition, the communities might be highly segregated within a ward. For example, in Bangalore upper caste neighborhoods are abutted by highly dense lower caste settlements. Even though the physical proximity might be less, the social distance

---

between these neighborhoods can be very high (A. Shaw, 2012). Clustering can happen even at a micro level – in one street, households belonging to one particular caste can reside and in the adjoining street, people belonging to a very different caste group might reside. Hence, the unit of analysis must be still smaller – say a street/census block if we want a clearer picture on caste based clustering based in a city.

For the first time, this study uses block level data to study residential segregation. A census enumeration block is roughly about 150-200 households which more or less constitutes a neighborhood. Hence a census block is optimum to study residential segregation. This study census-scale uses data collected by Government of Karnataka (2015) which provides block level data on caste for all 300 cities and towns of Karnataka. This chapter uses residential segregation as an indicator to test the hypothesis whether Urbanization weakens the structure of caste. This is tested by looking at caste based segregation in various cities ranked based on Zipf’s Law or the Rank-size Rule. If the hypothesis is true, then neighborhood level segregation should increase as the rank increases.

Data

Most of the studies on caste in India have used the nationally representative sample surveys such as the National Sample Survey (NSS), National Family Health Surveys (NFHS) and the India Human Development Survey (IHDS). These surveys except for IHDS do not ask about the individual caste, but major categories which are the Scheduled Castes (SCs), Scheduled Tribes (STs), other backward castes (OBCs) and the others. Even the all-India decennial census brackets individuals only into three major categories of SCs (Dalits), STs (Scheduled Tribes), and “others.” The last census which
recorded information on individual jatis (castes) was in 1931. We construct our diversity indices at sub-block level using household level caste data collected by the Government of Karnataka in 2015 (henceforth, GOKS).

The 1640 different caste categories enumerated by GOKS were recoded into 708 broad categories using detailed ethnographic and anthropological accounts (Anantha Krishna Iyer & Nanjundayya, 1928; Enthoven, 1990; Singh, 2002; Thurston & Rangachari, 1975). This recoding was done to primarily standardize jati names across the state and account for synonyms. For example, Agasa and Madiwala are synonyms for washerman caste. We also collapsed castes under all religions except the majority Hindu religion into single caste categories to better reflect actual patterns of spatial clustering and for reasons of data tractability. Dalit Christians are coded separately from all other Christians in the administrative classification and we retain this coding. The final 717 recoded jati categories were mapped to eight major administrative categories that are used the Government of Karnataka for purposes of affirmative action in election and selection processes of local governments, public employment, and admissions to institutions of higher education — “STs,” “SCs,” “I,” “II A,” “II B,” “III A,” “III B,” “others.”3 We use these eight administrative categories for our segregation analysis.

This data is unique in many respects. For the first time since the 1921 census, we have access to detailed household level caste data. Census does not reveal block level data for reasons of confidentiality and privacy. If we have to understand socio-economic

---

3 Category I mostly consists of nomadic and semi-nomadic castes which are not part of the caste hierarchy and avarna castes which are not part of the Scheduled caste list. Category I also consists of Scheduled Caste converts to Christianity. Category II A mostly consists of traditional occupational castes such as Agasa (washermen), Devanga (weaver), Kumbara (potter) etc. II B category includes Muslims. Dominant land holding communities and other land holding castes come under Category III. Category III A consists mainly of sub-castes of Vokkaligas, Reddys, and Kodavas. Category III B consists of sub-castes of Lingayats, Marathas, Jains, Bunts, Christians, and few other trading castes.
residential clustering, we need block level data which the national census does not provide. Calculation of diversity index depends on the proportion of a particular caste/class among caste/class groups in a specific area. This relative abundance (to use an ecological term) can only be calculated by surveying the entire community over the whole study area. Hence, sample surveys (though not impossible) cannot give a good assessment of clustering in a given area. We have considered 349 urban areas\(^4\) for our analysis.

**Methodology**

The literature on segregation has used a variety of different metrics to quantify segregation. The most commonly used metrics include the dissimilarity index (Duncan & Duncan, 1955; Jahn, Schmid, & Schrag, 1947; Taeuber & Taeuber, 1965), entropy class of metrics including the Theil index (Theil, 1972), and information theory based indices (Reardon & Firebaugh, 2002; Reardon & O’Sullivan, 2004). While each of these measures offer distinctive advantages, they also suffer from drawbacks that are particularly salient for our purposes (for example, problems with proportionate scaling the case of dissimilarity index, or the inability to discriminate between particular subgroup identities in the case of entropy metrics).

Figure 3 illustrates how extant indices are not well-suited for characterizing caste segregation in Indian cities. The figure depicts a city with three caste categories **RED**, **BLUE**, and **GREEN**. The city is divided into three wards (corresponding to the three rows

---

\(^4\) According to Census of India 2011, the definition of urban area is as follows;
1. All places with a municipality, corporation, cantonment board or notified town area committee, etc.
2. All other places which satisfied the following criteria:
   i) A minimum population of 5,000;
   ii) At least 75 per cent of the male main working population engaged in non-agricultural pursuits; and
   iii) A density of population of at least 400 persons per sq. km.
in the figure). Wards are segregated so that we have a RED ward, GREEN ward, and a BLUE ward. Each ward is further divided into two census blocks for a total of six census blocks in the city (the blocks are numbered A through F). In a city with nested aggregation structure (wards containing blocks in the present example), it is attractive to use a segregation metric that is additively decomposable. However, extant entropy metrics like the Theil index cannot discriminate between the three wards, or between the six census blocks. For example, census block B in the figure has the GREEN caste living in a RED dominated neighborhood; and block C has RED castes living in a GREEN dominated area. While both these blocks are equally segregated, they represent vastly different neighborhoods. Ideally a segregation metric must discriminate between these two blocks even while being additively decomposable so that segregation in the city can be decomposed into ‘within ward' and ‘between wards' components.

We use a recently developed divergence metric that achieves precisely this goal (Roberto, 2015). The divergence index, D is a relative entropy index that combines the desirable decomposition property of the entropy indices with the attractiveness of the widely used dissimilarity index. D is a norm-deviation metric that measures the ‘difference' between a normative benchmark distribution, Q and the empirical distribution P with N caste categories (Roberto, 2015):

\[
D(P \parallel Q) = \sum_{i=1}^{n} P_i ln \left( \frac{P_i}{Q_i} \right)
\]

In our analysis, we use the city-wide distribution of castes as the normative distribution (Q), and decompose segregation into between, and within ward components. To compute within-ward components, the ward distribution is used as the
normative reference. Additionally, we also compute divergence using ward as the reference.

Results

In this section, we present results from our investigation of residential segregation using eight caste categories as the axis of segregation. Figure 4 shows why the divergence index is an attractive metric to capture patterns of caste based residential segregation. The figure shows that there is no definitive relationship between divergence and diversity computed as ethno-linguistic fractionalization. The wards in Bengaluru, not surprisingly, coincides with the locus of the frontier. We use data from all 5481 urban wards in Karnataka. As discussed previously, the divergence index is a measure of how the ethnic composition in a given ward represents a deviation from the overall ethnic composition of the city of which the ward is a constitutive part. The divergence index is easy to interpret as a measure of relative segregation, especially when used in conjunction with the more traditional ethno-linguistic fractionalization index as a measure of social heterogeneity (the probability that any two randomly drawn individuals belong to different social groups). A high fractionalization index and a high divergence index indicated a ward that is heterogeneous but whose ethnic composition nonetheless is different from the city taken as a whole. There are several wards in the figure with high fractionalization but with low values of divergence indices. These points are from smaller cities that have very few wards for divergence to be significant. Out of the 355 urban centres in Karnataka, 137 have only a single ward. The median city in Karnataka has 16 wards and even more tellingly the 90th percentile for number of
wards in a city is 31, and there are only ten cities with more than 40 wards. This follows directly from the fact that city size distribution in Karnataka is power-law and consistent with a Zipf distribution as illustrated in Figure-1 and Figure-2. The fact that divergence cannot meaningfully be calculated for city-ward pairs for over 38% of the urban centres in Karnataka also points to the limitation of using ward as the primary unit of analysis in studies of urban spatial heterogeneity in India.

Figure 5, Figure 6, Figure 7 provide strong evidence for how extant analysis of residential segregation underestimates (or at least misrepresents) the actual patterns of segregation. Segregation within a ward is empirically as important as segregation between wards. The unit of analysis in segregation studies of urban India has always been the ward primarily because of data limitations. The lack of higher resolution ethnic composition data limits our understanding of the dynamics of how ethnic spaces and geographic spaces coevolve products of urbanization. In Figure 6, we present the kernel density plot of the “within” component of the divergence index. The overall divergence index is calculated at the census block level and decomposed into “within-ward” and “between-wards” components (recollect that the divergence index used here is perfectly additively decomposable). The divergence index at the block level is a good proxy for spatial segregation as it is a measure of the “difference” in the ethnic composition of individual census blocks and the city as a whole. A significant portion of the density graph is to the right of unity – suggesting that such wards are in cities where wards are similar to each other in ethnic composition so that the between component is actually negative with all the divergence coming from intra-ward variation. The corresponding negative “between” component is shown in Figure-7. Taken together, the density plots in
Figure 6 and Figure 7 demonstrate the need for spatial segregation studies in urban India to move beyond the ward as the preferred unit of analysis.

Figure 8 (Bengaluru) and Figure 9 (rest of urban Karnataka) provide preliminary evidence for how larger urban centers with more diverse populations do not necessarily result in reduction of residential segregation. LOESS smoothing for Bengaluru suggests that larger wards in Bengaluru are more heterogenous (as measured by fractionalization index) and have a smaller divergence index (larger wards resemble the city as a whole in terms of their ethnic composition). The positive relationship between divergence index and divergence in the case of Bangalore is also statistically significant (p < 0.01 in a bivariate linear model). However, this statistically strong positive relationship does not imply that spatial segregation necessarily reduces as the ward size increases (given aggregation effects). This is seen in Figure-9 (urban wards in Karnataka outside Bengaluru) where the sign of the relationship between ward rank and divergence index is reversed. Taken together, Figure 8 and Figure 9 suggest that there is no stable statistical relationship between divergence index and size of the ward.

In Figure 10 we present the relationship between divergence and city rank that further clarifies that there is no definitive relationship between urbanization intensity and spatial segregation. In particular, the bottom panel shows how city diversity (as measured by fractionalization index) is not significantly related to city rank. The lack of a monotonic relationship is further clarified in Figure 11 where we plot within-ward divergence against city rank. Neighborhood segregation is not statistically correlated with intensity of urbanization.

Our analysis suggests that using high resolution spatial data, and including actual caste identifiers rather than three broad aggregate categories reported by decennial
census data results in important modifications of extant segregation portraits. We have also demonstrated the usefulness of a new method for computing segregation. Divergence index is particularly well-suited for use with detailed caste data. In the next section we discuss the implications of the central result of our analysis – overall segregation being uncorrelated with degree of urbanization – for different social subgroups, and also discuss object lessons for research, policy, and praxis.

Discussion

Table-1 summarizes the patterns of extant spatial divergence and segregation. Table-1 seeks to answer the questions: how (if) does spatial segregation differently impact social groups? Are certain social groups more ghettoized? Are social groups differentially impacted by process of urbanization? Our analysis presented here provides first ever systematic census-scale evidence for differential impacts of urban segregation in India – in a discourse dominated by anecdotal accounts of ethnic space making. First, we computed mean block-level fractionalization for each of the eight social categories used in our analysis. The mean block level fractionalization presented in Table 1 represents the mean level of segregation experienced by each of the eight social groups. Muslims (administrative category 2B) live in most segregated urban blocks. A mean block fractionalization of 0.5 suggests that on an average Muslim households live in neighborhoods that are very homogenous. On an average, the probability that a randomly chosen household in the vicinity of a Muslim household will be a non-Muslim one is only 0.5. The second most segregated of social groups in urban Karnataka is the SC group. The average SC household lives in a neighborhood with a block fractionalization of only 0.6. Table 1 also shows the value of using high-resolution census
block level data – especially to understand the differential patterns of spatial segregation across social groups. For example the ward level fractionalization experienced by the average SC household is not very different from that of an upper caste Hindu household (category OTH) but at the block level SC households live in significantly more homogeneous neighborhoods than upper caste Hindu households. This is also corroborated by the fact that the mean divergence indices measuring how individual blocks in a ward are different from the ward as a whole are higher for Muslim and SC households.

The means reported in Table 1 obfuscate even more significant differences in spatial divergence experienced by various social groups across the distribution. Figure 13 shows distribution of block level fractionalization experiences by households belonging to each one of the eight social groups used to calculate fractionalization. The distribution of Muslim households (category 2B) is distinctively bimodal. The bimodal distribution suggests that while a significant number of Muslim households live in highly segregated ethnic ghettos, there is a substantive part of the Muslim distribution coincides with distributions for other social groups. The distribution for SC is also bimodal though not as pronounced as the one for Muslim households. Figure 14 presents ward-level rather than block-level fractionalization and this analysis once again underscores the utility of using high resolution data. The ward level distribution for Muslim households is multi-modal, and consistent with findings reported in Table-1. Figure 15 that reports distributions for divergence index provides the clearest evidence for why intra-ward segregation is especially significant for marginalized social groups.

While our analysis provides the most detailed portrait of spatial segregation in urban India, our snapshot provides little insights into the causal pathways that can
explain the linkages between urbanization and ethnic space making. Our detailed snapshot from high resolution ethnic data and high resolution spatial data has identified three spatial divergence patterns that all call for detailed micro-level ethnographic studies to uncover the long-term processes that generate these patterns. First, we have provided conclusive evidence for how spatial segregation is largely independent of level of urbanization. This result poses a significant challenge to one of the bedrock normative promises of urbanization in India. As we have noted in the introduction to our analysis, one of the key sources of normative support for urbanization – especially among the most marginalized social groups – has been the possibility that urbanization can help in remaking ethnically segregated physical spaces in an agrarian regime. For scholarship, policy, and praxis to grapple with this conundrum, we need a more acute understanding of the actual processes that result in replication of ethnically segregated spaces in urban centres. Second, we have shown how urbanization has a differential impact on various social groups – marginalized social groups occupy more segregated spaces than socially dominant groups. Each one of these two patterns of ethnic space making require detailed and widespread ethnographic investigations to understand the causal pathways. Another significant limitation of the portrait we have presented here is that we have not been able to control for non-ethnic characteristics of households. For example, we do not know if the pattern of social segregation of space that we have presented here holds across economics classes.\(^5\) Third, and the most significant policy relevant results from the analysis presented here is the crying need for spatial divergence studies to look beyond urban wards as the unit of analysis. While data limitations will continue to be a binding

\(^5\) Though GOKS collected data on class markers (self-reported income and educational attainment), this information is not yet available for analysis. We will revisit our analysis presented here when such information becomes available.
constraint, our analysis suggests the need for carefully constructed sample surveys that can delineate intra-ward spatial divergence.

Figure 1 Rank-Size Distribution of Cities. The chart shows the rank-size relationship for all cities in our dataset. The blue colored point is Bengaluru (a city of over ten million residents), over nine times the size of the second ranked city. The figure shows how the city size distribution in our dataset is consistent with so-called Zipf-law. In the Indian context, given the census definition of an urban settlement, city rank also serves a proxy for intensity of urbanization.
Figure 2 Rank-Size Distribution of Wards. Not surprisingly, wards in Bengaluru (in blue) are larger than wards in every other city save isolated instances from second and third ranked cities. Given the Zipf-law like distribution of wards, we use ward rank as a proxy for level of urbanization (also cf. notes to Figure -1).
Figure 3 Discriminating Ethnic Fractionalization in a City with Three Caste Groups (Red, Blue, and Green). The figure shows three wards (the three rows of the figure) and six census blocks (numbered A through F). Each of the blocks have the same ethnic fractionalization but from the perspective of caste based spatial clustering each of the blocks are distinct. See main text for details.
Figure 4 Fractionalization and Divergence. The chart illustrates the differences between concepts of ethnic fractionalization and divergence index that we use to measure urban segregation in this paper. The figure represents diversity in all urban wards in Karnataka (n = 5481). Fractionalization is computed using the eight-fold administrative taxonomy, and divergence is computed as the “difference” in ethnic compositions of wards and the city. The Blue dots are from wards in the largest city in our dataset, Bengaluru – to city with more than ten million residents. The red dots are from wards in other cities and towns in the dataset. Divergence index represented here measures how census blocks within a ward are different from ward as a whole.
Figure 5: Blocks are Different from Wards. Panel-A shows the kernel density plots (blue is Bengaluru and pink is rest of urban Karnataka) for fractionalization index within a ward. Panel-B plots these distributions for divergence index (measuring how blocks within a ward are different from the ward). $n = 5481$ wards. Taken together, these density plots provide evidence for how ward level aggregate data understate the level of segregation in Indian cities. See main text for more explanation.
Figure 6: Decomposing Divergence Index. The figure presents the kernel density plot (gaussian fit) on the distribution of the within-ward component of total divergence. The divergence between individual census blocks and the city are decomposed as “within-ward” and “between-wards” components. The second mode in the density plot corresponds to smaller towns in the dataset that have only a single ward (hence all divergence is “within” component). A “within” component greater than unity indicates that wards in these cities are closer to each other in population characteristic than blocks within a ward are. This density plot underscores the need for going beyond ward level segregation analysis. See text for more explanation (also confer Figure 7).
Figure 7: Decomposing Divergence Index. The figure presents the kernel density plot (guassian fit) on the distribution of the between-ward component of total divergence. The divergence between individual census blocks and the city are decomposed as “within-ward” and “between-wards” components. The first mode in the density plot corresponds to smaller towns in the dataset that have only a single ward (hence all divergence is “within” component). A “between” component greater than zero indicates that wards in these cities are closer to each other in population characteristic than blocks within a ward are. This density plot underscores the need for going beyond ward level segregation analysis. See text for more explanation (also confer Figure 6A).
Figure 8 Segregation and Ward Size, Bengaluru (n= 198). Panel-A shows that there is a (weak) negative relationship between ward-rank and fractionalization. Larger wards, not surprisingly are more diverse. Panel-B shows that larger wards also contain less segregated neighborhoods (as measured by divergence index). See text for more explanation.
Figure 9  Segregation and Ward Size (all cities except Bengaluru, n = 5283). The left panel plots ward-level fractionalization for all urban wards outside Bengaluru against ward rank (ranked by population size). Wards are diverse across the spectrum as indicated by a lack of strong relationship between rank and fractionalization. The right panel however points to spatial patterns of segregation within the wards. The right panel plots divergence computed at the ward level – an indicator of how census blocks within a ward are different in their ethnic composition from the ward as a whole. See main text for more explanation.
Figure 10  City Size and Spatial Divergence. The top panel shows divergence index for all census recognised urban centres in Karnataka (n = 355). The divergence index is an indicator of how individual census blocks are ‘different’ from the city or town as a whole in terms of their ethnic composition. The bottom panel simply plots ethnic fractionalization for each of these cities. See main text for more details.
Figure 11 Within-Ward Divergence and City Size. The figure plot within-ward divergence (an indicator of census blocks within a ward are different in their ethnic composition from the ward as a whole. The figure shows that intra-ward spatial segregation displays a very weak relationship in terms of how it varies with city size (n = 355). See main text for more explanation.
Figure 12 Mean Block-Ward Divergence. Each of the eight panels show mean divergence of ethnic distribution between blocks and wards in each of the 355 urban centres (ranked by their population). The divergence index for each subgroup is a weighted mean (weighted by block population) of block-ward divergence for households in the subgroup. The figure suggests that there is no monotonic relationship between intensity of urbanization (city rank) and patterns of spatial divergence. See main text for further explanation.
Table 1: Patterns of Spatial Divergence in Urban Karnataka, 2015. The table summarizes patterns of spatial divergence presented in this paper. We use the eight-fold administrative social category used by government of Karnataka in our analysis. Constituent Jati categories for each of this eight subgroups are described in column-2. The third column records the percentage of households in each subgroup that resided in census designated urban centres in 2015 (this data is computed from census-scale household data, n = 13,255,421). The fourth column records the mean block-level ethnic fractionalization index for each of the eight sub-groups (weighted by block population). The table shows that Muslims and Scheduled Caste groups are likely to live in more homogenous neighborhoods. The next column presents the same information at the ward level. The difference in ethnic fractionalization values at block and ward levels shows how urban wards are heterogeneous and spatial segregation is best studied at the census block level. The last column records the mean divergence index (a measure of how blocks within a ward are different from the ward in terms of its ethnic composition). See main text for further details.

<table>
<thead>
<tr>
<th>Administrative Caste Category</th>
<th>Description</th>
<th>% of urbanized households (within the group)</th>
<th>Mean Block Fractionalization (Weighted)</th>
<th>Mean Ward Fractionalization (Weighted)</th>
<th>Mean Divergence Index, Block-Ward (Weighted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nomadic and semi-nomadic castes</td>
<td>24.03</td>
<td>0.70</td>
<td>0.79</td>
<td>0.27</td>
</tr>
<tr>
<td>2A</td>
<td>Traditional occupational castes</td>
<td>25.95</td>
<td>0.69</td>
<td>0.79</td>
<td>0.26</td>
</tr>
<tr>
<td>2B</td>
<td>Muslims</td>
<td>38.46</td>
<td>0.50</td>
<td>0.69</td>
<td>0.32</td>
</tr>
<tr>
<td>3A</td>
<td>Mainly Vokkaligas, Reddys, and Kodavas</td>
<td>26.26</td>
<td>0.72</td>
<td>0.79</td>
<td>0.25</td>
</tr>
<tr>
<td>3B</td>
<td>Mainly Lingayats, Marathas, Jains, Bunts, Christians</td>
<td>24.68</td>
<td>0.70</td>
<td>0.79</td>
<td>0.26</td>
</tr>
<tr>
<td>OTH</td>
<td>Brahmins and other castes</td>
<td>43.42</td>
<td>0.70</td>
<td>0.80</td>
<td>0.27</td>
</tr>
<tr>
<td>SC</td>
<td>Scheduled Castes</td>
<td>22.33</td>
<td>0.60</td>
<td>0.78</td>
<td>0.30</td>
</tr>
<tr>
<td>ST</td>
<td>Scheduled Tribes</td>
<td>15.06</td>
<td>0.68</td>
<td>0.79</td>
<td>0.27</td>
</tr>
</tbody>
</table>
Figure 13: Distribution of Block Fractionalization Index by Social Category. Distributions presented here are for approximately five million urban households across all 355 urban centres in Karnataka. The Block Fractionalization Index is the fractionalization index calculated for all urban census blocks in Karnataka using the eight-fold administrative taxonomy of social groups. The distinctive bi-modal distribution for Muslim and SC households (categories 2B, SC) points to why mean divergence reported in Table-1 must be interpreted with care. See main text for more explanation.
Figure 14: Distribution of Ward Fractionalization Index by Social Category. Distributions presented here are for approximately five million urban households across all 355 urban centres in Karnataka. The Ward Fractionalization Index is simply the fractionalization index calculated for all urban wards in Karnataka using the eight-fold administrative taxonomy of social groups. The thick left tail for Muslim households (category 2B) points to why mean divergence reported in Table-1 must be interpreted with care. See main text for more explanation.
Figure 15: Distribution of Ward-Block Divergence Index by Social Category. Distributions presented here are for approximately five million urban households across all 355 urban centres in Karnataka. The Ward-Block Divergence Index measures how the ethnic composition of the ward is different from the census blocks within a ward. The multi modal density maps for Muslim households (category 2B) points to why mean divergence reported in Table-1 must be interpreted with care. See main text for more explanation.
References


Miklian, J., & Sahoo, N. (2016). Supporting a More Inclusive and Responsive Urban India. Available at SSRN.


Omvedt, G. (1978). Class Struggle or Caste War’. *Frontier, Autumn No, 2*.


Whyte, W. F. (1943). Street corner society; the social structure of an Italian slum.


<table>
<thead>
<tr>
<th>WP No</th>
<th>Title</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019-04</td>
<td>Village in the City: Residential Segregation in Urbanizing India</td>
<td>Bharathi N., Malghan D., Rahman A.</td>
</tr>
<tr>
<td>2019-03</td>
<td>Inequality in a Global Perspective</td>
<td>Kanbur R.</td>
</tr>
<tr>
<td>2019-02</td>
<td>Impacts of Minimum Wage Increases in the U.S. Retail Sector: Full-time versus Part Time Employment</td>
<td>Yonezawa K., Gomez M., McLaughlin M.,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018-14</td>
<td>Commercialization of a Demand Enhancing Innovation by a Public University</td>
<td>Akhundjanov, S. B., Gallardo, K.,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>McCluskey, J., Rickard, B</td>
</tr>
<tr>
<td>2018-12</td>
<td>Sustainable Development Goals and Measurement of Economic and Social Progress</td>
<td>Kanbur, R., Patel, E, and Stiglitz, J.</td>
</tr>
<tr>
<td>2018-10</td>
<td>Inequality and the Openness of Borders</td>
<td>Kanbur, R.</td>
</tr>
</tbody>
</table>

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.