Housing market and employment discrimination in the Mumbai Metropolitan Region - a spatial analysis of caste-based segregation

Aaditya Warrier

1. Introduction

Mumbai, India's financial and commercial capital, is both its wealthiest city and its most diverse. The former headquarters of the Bombay Presidency during British rule, the city had long been home to large minority communities from around the province, but was always dominated by Maharashtrians and Gujaratis, the ethnic natives of the region. Post-independence, however, the city was (and continues to be) shaped by waves of immigration from across the state and the country, exponentially growing its population and consequently its urbanized area to accommodate these new residents. Large swathes of land to the north of the city, previously independent towns or villages, were annexed by administration. Mumbai is no stranger to expansion - indeed, the original island city owes its mere existence to land reclamation by the British - but its scale and scope in the past few decades has been unprecedented. In 1951, the city covered 231.5 km^2 and housed 2.3 million people. By 2011, the city's metropolitan region (henceforth shortened as MMR) spanned 6,328 km^2 with 22.8 million people, with the city proper housing 12.2 million over a mere 603 km^2. Today, Maharashtrians and Gujaratis only make up 42% and 19% of the city's population respectively.

These rapid changes in demography have brought with them a significant rise in ethnic, religious, and caste tension. Combined with rising real estate prices - the highest in India and in the top 20 globally - these have manifested in *severe* social segregation in the housing market. Despite recent laws that have attempted to curb the practice (Koppikar 2023), many apartment buildings have been found outright barring Muslims, lower caste individuals, and specific ethnicities; failing that, they ban "non-vegetarians" in order to circumvent the law, as lower/middle caste Hindus and religious minorities alike are more likely to be non-vegetarian. (T. Staff, n.d.). From Gujarati-preferred colonies in suburban neighborhoods like Vile Parle and Ghatkopar to the ghetoization of Muslims in Mumbra, many in the segregated lives.

When not motivated by self-preservation as a social minority - as is the case for Christians strangehold on housing in Bandra, for example - the common thread uniting these blatantly discriminatory acts is likely a strong **casteist** sentiment. With housing societies going as far as having separate elevators for residents and "servants" (2020), we see a strong revulsion of living in, or even sharing any space, as people with perceived undesirable characteristics, linked to the idea of the "impurity" of lower castes. Even when ethnic solidarity has been observed amidst rising tensions, caste fractures have run high Mumbai; for example, the Shiv Sena, a powerful, right-wing, pro-nativist political party that came into power a few decades ago rallying under preferential treatment for Maharashtrians, covertly opposed affirmative action for lower caste Maharashtrians (see Worli riots, opposition to Mandal Commission). This discrimination can extend to "natives" of the region as well - numerous stories have emerged in recent years (M. Staff 2021) of Maharshtrians being denied housing by Gujaratis (on average, local Gujarati communities are wealthier and of higher caste than Maharashtrians).

This project hence aims to explore to what extent Dalits, Adivasis, and other lower caste groups have been pushed to the margins of the city, as well as how current demographics of the region reflect the characteristics of immigrating populations. The primary objective of the project is descriptive rather than inferential or predictive, and aims to obtain a spatially-informed understanding of segregation both within Mumbai's exurbs and between the city and its fringes. Spatial segregation in Mumbai's housing has been well-established in the literature (Shaban and Aboli 2021); however, little analysis has been conducted on segregation in the wider metropolitan region, outside the boundaries of the city proper. In a city with limited land space, segregation effectively incentivizes those shut out of its restrictive market to either live in slum-like informal tenements or to seek housing in far-flung satellite cities in the MMR. Not coincidentally, real estate prices also tend to decrease as one moves northwards and east-wards from the old city, further pushing lower-income individuals (who are disproportionately lower-caste and/or Muslim) outwards. This has the potential for severe business impacts (see Bharathi, Malghan, and Rahman 2023) - daily commutes of three hours or more are not un-common in the region, as a majority of employment tends to be concentrated in industrial areas and business districts in the city proper, even today. There is also evidence that these exurbs are less developed and have reduced access to public services(pre-print from Asher et al. 2023), further underlining the necessity of this study.

2. Data

The data used in this study comes from The Socioeconomic High-resolution Rural-Urban Geographic Platform for India Project (abbrv. SHRUG). SHRUG is an open access repository that aims to link hundreds of Indian datasets comprising socio-economic surveys, censuses, and environmental studies with common geographic identifiers. Owing to the nation's size and diversity, Indian record-keeping is notoriously inconsistent, with formats, naming conventions, and even languages often differing within the same state. SHRUG obtains a level of standardization and geographic granularity that cannot be found in any single government dataset -

by splitting the nation into 576,153 uniform areal units termed "shrids" and reaggregating all variables to these, it provides a largely consistent framework facilitating analysis at a deeper level than states, districts and sub-districts. This is the primary reason that this project utilizes **2011 Population Census Abstract (PCA), Socio-Economic and Caste Survey (SECC), and Economic Cenus (EC) Data** (Registrar General and Census Commissioner 2011; Ministry of Rural Development 2011; Asher et al. 2021) over more recent data like the National Family Health Survey from 2019-21 - cleaning the latter to a level acceptable for analysis would have been a months-long project, as it is not currently available in SHRUG.

The 2011 PCA, EC, and SECC datasets are the most current census data available for India; although usually conducted decennially, the 2021 census was postponed due to COVID and other political reasons. As stored at the shrid and subdistrict level in SHRUG, they contain a multitude of variables on employment, occupation, literacy, property ownership, and consumption, as well as demographic information like gender, caste, income, and education status. Key variables of interest were the Scheduled Caste/Scheduled Tribe (SC/ST) population percentages, home/land ownership shares, and employment/business ownership by different castes. Note that SC/ST are government designations for Dalits and Adivasis.

Surprisingly missing from the dataset was a breakdown of religions and ethnicities in each shrid. This could have provided additional nuance to our analysis, especially when looking at how they intersect with caste; contrary to popular belief, many non-Hindu Indians also have caste distinctions, while an ethnic breakdown could contextualize discrimination in terms of the familiarity of individuals with specific sub-castes not present in their native regions.

2.1 Data Cleaning

As SHRUG stores district/sub-district data separately from shrid-level data, with no linking variables between the two, finding the specific shrids that comprised the MMR outside of the city proper proved a challenge. This was compounded by the fact that the MMR is defined by the Municipal Corporation and Councils that comprise it, the exact boundaries for which occasionally span multiple state-level districts or include very small portions of specific sub-districts. To rectify this problem as much as possible, the spatial intersection (using the st_interescts predicate) between the most relevant sub-districts of the MMR and all shrids in the PCA data was computed. This, however, missed out on a few areas of interest, most notably the majority of the municipal council of Khopoli, while including a few parcels of rural land that were not relevant. The cleaned dataset contained 1273 shrids with a population totaling roughly 22.1 million - this was deemed an acceptable enough difference from the true population of 22.8 million.

Another issue encountered was that the city proper was only comprised of two shrids, each spanning the Mumbai City and Mumbai Suburban districts respectively. The authors of SHRUG note this as an issue in their documentation - specifically, that breaking up census data into shrids was challenging in and around major cities, given the occasional lack of detail

in the census - and this hampered some of the comparative analysis that was planned with regards to segregation in the city vs. the rest of the MMR.

Finally, there were some missing shrids in the EC and SECC datasets as compared to the PCA dataset; specifically, 42 shrids were missing from SECC and another 60 from EC. As a result, only PCA variables are available for those vs. the full set of PCA, EC SECC variables for the rest.

2.2 Exploration

Table 1: Caste demography of the MMR vs. India

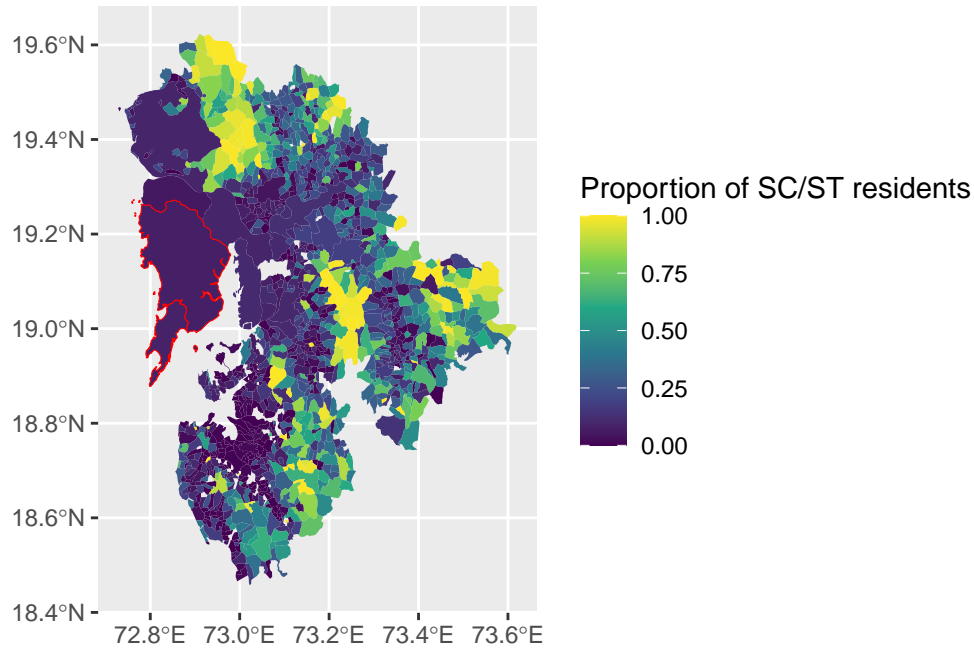|	Region	SC Population (% of total)	ST Population (% of total)	Combined SC/ST Population (% of total)
1	Mumbai (city proper)	6.46	1.04	7.50
2	MMR	6.76	3.13	9.89
3	MMR (excluding city proper)	7.12	5.56	12.68
4	India	16.60	8.60	25.20

Table 1 displays a breakdown of caste percentages

4

This is confirmed by Figure 1, which displays a chloropleth-map of all regions of the MMR (including Mumbai City/Suburban districts). As before, note that all of the city proper is are two large shrids, outlined in red on the map. We observe a pattern of increasing SC/ST percentages as we move northwards from the original island city

with high percentages observed in the Kalyan and Bhiwandi districts as well as parts of Vasai and Virar. A similar trend is observed to the east in Karjat and Panvel as well as the souhermost parts of Alibag/Pen district.

This reinforces initial speculation as to the pricing out

This indicates

Geary's C	P-value
0.543	0

Conducting a hypothesis test for Geary's C shows us clear evidence of positive spatial auto-correlation, indicating that a shrid with a high SC/ST proportion is likely to neighbor a shrid with a very similar proportion. These set the context

3. Spatial Analysis

3.1 Methodology

3.1.1 Primary Outcome

The primary outcome of this study is a set of descriptive statistics to describe spatial, caste/tribe-based segregation at a shrid-level.

is the global spatial dissimilarity index, as formulated in [@], of SC/ST individuals vs. all other individuals. A spatial analog to the dissimilarity index

will be tested across different bandwidths

This index was chosen for its prior use in the Indian context (see Haque et al. 2021, which explores spatial segregation in Kolkata and Bengaluru) as well as its ability to mitigate the so-called "checkerboard problem" (Reardon and O'Sullivan 2004), of great importance in a city like Mumbai where expansions of its territory and local geography have been shaped significantly by different waves of immigration.

These metrics were calculated using the *seg* package in R, developed in part by O'Sullivan. As the *seg* package has been temporarily unavailable on CRAN since Feb 2024 due to dependency issues, an older version of the package available on the archive was used instead; the tar.gz can be found in the project repo.

Significance will be tested using a pseudo-significance "p-value" measure, computed on random permutation tests of dissimilarity across different spatial arrangements of the region (using the procedure described in Feitosa et al. 2007, 309–10, which was computed on a very similar index).

3.1.2 Secondary Outcomes

The secondary outcome of this study is the same set of spatial statistics, but computed at a subdistrict level.

Tests if the city is split more along income lines than caste, acknowledging the high correlation between the two

Test at subdistrict level

Avoid the "grid" problem - use subdistrict breakup as well??

3.2. Results

Table 2: Secondary Outcome Results

SC	ST	Combined SC/ST
0.102	0.428	0.154

```
# A tibble: 0 x 0
```

4. Conclusion and Discussion

The limitations of this study arose primarily from data availability and quality. Ideally, this study would have been conducted at the Ward Level

Methodologically, we must also acknowledge that many of these measures are, relatively speaking, new in the literature and each suffers from limitations unique to them. For example, even though Hong's measure meets compositional invariance criteria, it suffers from MAUP issues, something that may be worsened by the alternatingly arbitrary and non-arbitrary basis on which shrids are defined in SHRUG.

References

2020. *IndiaTimes*. https://www.indiatimes.com/trending/social-relevance/society-in-mumbais-bandra-is-allegedly-using-separate-lifts-for-owners-and-servants-511234.html.

Asher, Sam, Kritarth Jha, Paul Novosad, Anjali Adukia, and Brandon Joel Tan. 2023. "Residential Segregation and Unequal Access to Local Public Services: Evidence from 1.5m Neighborhoods in India∗." In. https://api.semanticscholar.org/CorpusID:260510509.

Asher, Sam, Tobias Lunt, Ryu Matsuura, and Paul Novosad. 2021. "Development Research at High Geographic Resolution: An Analysis of Night-Lights, Firms, and Poverty in India Using the Shrug Open Data Platform." *The World Bank Economic Review* 35 (4).

Bharathi, Naveen, Deepak Malghan, and Andaleeb Rahman. 2023. "Ethnic Diversity and Economic Development with Spatial Segregation." *Economics Letters* 222: 110951. https://doi.org/https://doi.org/10.1016/j.econlet.2022.110951.

Feitosa, Flávia, Gilberto Câmara, Antonio Monteiro, Thomas Koschitzki, and Marcelino Silva. 2007. "Global and Local Spatial Indices of Urban Segregation." *International Journal of Geographical Information Science* 21 (March): 299–323. https://doi.org/10.1080/13658810600911903.

Haque, Ismail, Dipendra Nath Das, Priyank Pravin Patel, and Md Hasnine. 2021. "Towards an Enhanced Understanding of Caste-Based Residential Segregation in Indian Cities: Reflections from Kolkata and Bengaluru." *Spatial Demography* 9 (2): 187–211. https://doi.org/10.1007/s40980-021-00085-8.

Koppikar, Smruti. 2023. "Mumbai Naama: Politicians and Law Do Not Have Answers to Mumbai's Housing Discrimination." *Free Press Journal*. Free Press Journal. https://www.freepressjournal.in/analysis/mumbai-naama-politicians-and-law-do-not-have-answers-to-mumbais-housing-discrimination.

Ministry of Rural Development, Government of India. 2011. "Socio Economic and Caste Census."

Reardon, Sean F., and David O'Sullivan. 2004. "Measures of Spatial Segregation." *Sociological Methodology* 34: 121–62. http://www.jstor.org/stable/3649372.

Registrar General, Office of the, and India Census Commissioner. 2011. "Census of India."

Shaban, Abdul, and Zinat Aboli. 2021. "Socio-Spatial Segregation and Exclusion in Mumbai." In *Urban Socio-Economic Segregation and Income Inequality: A Global Perspective*, edited by Maarten van Ham, Tiit Tammaru, Rūta Ubarevičienė, and Heleen Janssen, 153–70. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-64569-4_8.

Staff, Maktoob. 2021. "Mumbai: No Flat for Marathi-Speaking Man in 'Gujaratis Only' Building, Owners Arrested." *Maktoob Media*. https://maktoobmedia.com/india/mumbai-no-flat-for-marathi-speaking-man-in-gujaratis-only-building-owners-arrested/.

Staff, TOI. n.d. "Debarring from HSG Societies on Basis of Faith, Food Goes Against Spirit of Co-Op Law: Mumbai News - Times of India." *The Times of India*. TOI. https://timesofindia.indiatimes.com/city/mumbai/debarring-from-hsg-societies-on-basis-of-faith-food-goes-against-spirit-of-co-op-law/articleshow/104056246.cms.

Appendix