Residential Segregation and Unequal Access to Local Public Services in India: Evidence from 1.5m Neighborhoods*

Anjali Adukia† Sam Asher‡ Kritarth Jha§ Paul Novosad¶ Brandon Tan‖

July 2022

Abstract

There is little quantitative evidence on the role that neighborhood settlement patterns play in mediating inequality in rapidly urbanizing lower-income countries. This paper helps to close this gap by analyzing settlement patterns, access to public services, and economic outcomes across 1.5 million neighborhoods for two of India's marginalized communities: Scheduled Castes (SCs) and Muslims. Patterns of segregation and unequal access to public services in India's cities largely replicate the striking inequalities in its villages. Segregation of SCs and Muslims in both cities and rural areas is comparable in magnitude to that of Black people in the U.S.. We find that public schools and hospitals are systematically located *away* from neighborhoods where Muslims and SCs are concentrated. These disparities are driven almost entirely by differences across neighborhoods and within towns, and are thus hidden by the more aggregated data typically used to study inequality. Inequality in public service is thus driven by local decision-making — the most informal and least systematically observed level of government. Children from all groups who grow up in urban minority neighborhoods attain less schooling, even after controlling for parent education, and household and neighborhood consumption. Unequal access to public facilities in India's highly segregated neighborhoods may be a significant contributor to disadvantages faced by marginalized groups.

*We are grateful for stellar research assistance from Ryu Matsuura and Alison Campion, and for helpful discussions in seminars at the Center for Policy Research, Centre de Sciences Humaines, IDinsight, Harvard Urban Development Workshop, IZA Labor Markets in South Asia Conference, NBER SI DEV Pre-Conference Workshop, Regional Disparities Seminar, Delhi Political Economy Workshop, SAIS, and American University.

†University of Chicago and NBER, adukia@uchicago.edu

‡Johns Hopkins SAIS, sasher2@jhu.edu

§Development Data Lab, jha@devdatalab.org

¶Dartmouth College, paul.novosad@dartmouth.edu

‖Harvard University, btan@g.harvard.edu

1 Introduction

Residential neighborhoods are widely understood to be key determinants of upward mobility, human capital, access to employment, and other long-run socioeconomic outcomes (Kling et al., 2007; Chetty et al., 2016; Chetty and Hendren, 2018; ?). The concentration of marginalized social groups into poorer neighborhoods is understood to be a major driver of the intergenerational persistence of cross-group inequality in many contexts (Ananat and Washington, 2009; Alesina et al., 2011; Boustan, 2013; Cutler et al., 2008). Residential segregation can have a range of deleterious consequences: members of segregated groups may face worse discrimination in terms of provision of public services, they may have worse access to employment and employment networks, and stereotypes in the wider population may be more difficult to break, among others (Cutler and Glaeser, 1997). Because residential settlement patterns tend to be highly persistent, these disadvantages may be particularly difficult to address.

Most of the empirical literature on residential segregation and neighborhood effects comes from developed countries, in large part due to the paucity of cross-neighborhood data in developing countries. However, the role of neighborhoods is particularly important to study in poorer countries. Cleavages across social groups are just as important, if not moreso, in developing countries as they are in developed countries. However, these countries are rapidly urbanizing, and thus the scope for policy to affect urban settlement patterns (which may stay in place for decades) is much greater than in richer countries. The extent to which cross-group disparities will become entrenched by urban settlement patterns remains to some extent a policy choice in the cities of the world which are still rapidly growing.

Studying outcomes and service delivery at the neighborhood level is also important because the neighborhood is the level at which many public services are accessed, even if it is not the level at which their allocation is most closely studied. In India, the focus of this paper, federal and state policies largely prescribe the allocation of services at aggregate levels (state, district, or subdistrict), while the neighborhood-level provision of those services is determined through less formal local processes. This creates a blind spot for policy makers. Policy makers aiming to equalize access to public services, but who can only observe the world at an aggregate level (for both allocation and evaluation), may fail to achieve their progressive objectives if cross-neighborhood disparities differ markedly from cross-region disparities.

For example, the Indian government has frequently targeted extra public services to districts with many Scheduled Castes. A district-level analysis might further suggest that disparities in access to services have closed or even reversed. But if those services are delivered entirely to majority neighborhoods (foreshadowing some of our results), the cross-neighborhood allocation could entirely undermine the original policy.

In this paper, we study the relationship between neighborhood-level settlement patterns and access to public services across 1.5 million urban and rural neighborhoods spanning all of India, focusing on outcomes for Muslims and members of Scheduled Castes. India is an interesting context in which to study these questions for several reasons. First, it is huge: the marginalized groups that we study number over 300 million individuals. Second, disparities across these groups are rooted in historical inequalities that have persisted for generations, but the extent to which those inequalities are being changed by market liberalization and urbanization remains an open question. Third, the policy and planning process in India remains focused on disparities at aggregate levels like the district; a recognition of how these aggregate plans translate to neighborhood-level outcomes is essential to understanding whether these policies are achieving their objectives.

We have three primary aims. First, we document the extent of residential segregation in rural and urban areas. Second, we describe how three public services—primary schools, secondary schools, and medical clinics—are distributed across marginalized group (MG) and non-marginalized-group (non-MG) neighborhoods. Third, we study the educational outcomes of children who live in MG neighborhoods, providing suggestive evidence on the impacts of residential segregation.

Our primary data sources are a pair of Indian administrative censuses that used the same coding scheme for small neighborhoods — about 100–200 households, or 500–1000 individuals each. The Socioeconomic and Caste Census (SECC 2012) is an asset census, which describes a short list of assets at the household level for *every household in India*, as well as a household roster with education, occupation, and SC status for every household member. Crucially, early listings of the SECC were released with respondent names; the distinctive naming patterns of Muslims allow us to predict the religious identity of household members (as Muslim or non-Muslim) with high accuracy.[1] Our second data source, the Indian Economic Census, records data on India's 65 million non-farm establishments. It is conventionally used to study firms, but it also records data on public schools, health clinics, and hospitals, making it the only large-sample data source (to our knowledge) that can identify the availability of public services at the neighborhood level. Combining these datasets, we can document individual demographics, socioeconomic outcomes, and neighborhood-level public services in 418,687 urban neighborhoods (in 3579 cities and towns) and 1,166,049 rural neighborhoods (in 420,000 villages).[2]

SCs and Muslims make up similar population shares in the country (17% and 14% respectively in 2011), but the histories of these two groups are distinct. Scheduled Castes have been historically consigned to the lowest occupational rungs of society for over a thousand years, but have been targeted by decades of affirmative action policies since independence; some studies suggest these programs have had positive effects (Asher et al., 2021a; Gulzar et al., 2020).

Different Muslim groups have historically occupied heterogeneous positions in Indian society over the generations; some Muslims are descendants of India's 15th to 18th century ruling classes, while others descend from lower-caste groups who converted to Islam to escape their status at the bottom of the social pyramid. Groups from both of these heritages increasingly find themselves politically marginalized and threatened. A large literature has discussed the relative outcomes of Scheduled Castes, and there is a more moderate literature on Muslims.[3] While disparities in access to public services have been documented for both of these groups, there is little prior work

[1] We classify names as Muslim or non-Muslim using a long short-term memory neural network based on a training set of two million takers of the Indian Railways Exam. The out-of-sample accuracy against a set of manually classified names was 97% (Ash et al., 2022).

[2] Due to incomplete data in the SECC, this represents about 80% of rural subdistricts and 45% of cities and towns. The excluded subdistricts and towns are broadly similar on a wide range of variables to those in the study.

[3] On Muslims, see, for example, Basant and Shariff (2010) and Jaffrelot and Gayer (2012).

on disparities at the neighborhood level—the level at which these services are typically accessed.[4]

We present three key findings. First, we find that Muslims and SCs are have highly segregated residence patterns.[5] On average, Muslims are less segregated than U.S. Black people, while SCs are more so. However, the distribution of Muslim shares across neighborhoods is highly bimodal; even though Muslims are less segregated than SCs on average, a greater share of Muslims live in neighborhoods that are almost entirely Muslim. [6]

Urban segregation is only slightly lower than rural, and urban and rural segregation are highly correlated across regions, for both Muslims and SCs. Indian cities are replicating the social group settlement patterns that have been in place for hundreds of years in villages.[7] Highly segregated cities are on average poorer than less segregated cities; this pattern holds for all social groups, a notable difference from the U.S. context where segregated cities are primarily associated with worse outcomes for Black Americans. Younger cities are less segregated, but the differences are very small, suggesting that the social and economic patterns that shaped past residential choices in cities are strongly persistent.

Second, we describe the relationship between MG share and public good availability at the neighborhood level, controlling for city fixed effects; the result tells us how a given stock of public facilities is allocated within cities (and within rural subdistricts or villages). We find large and systematic disadvantages in access to public facilities—public primary schools, secondary schools, and health clinics—in neighborhoods where marginalized groups are concentrated. Access to public facilities in Muslim neighborhoods is universally worse than in SC neighborhoods. This result echoes a consistent finding across the qualitative literature, where Muslims report difficulty in getting public service delivery from their representatives (Jaffrelot and Gayer, 2012).

[4]Because the Population Census records Scheduled Caste shares and the presence of a set of public services at the village level, the relationship between the village-level rural SC share and public services has been studied (Banerjee and Somanathan, 2007). However, village or neighborhood-level access to public services has not been studied on a national scale for Muslims, nor urban access to local services for either group.

[5]We primarily use the dissimilarity index, but results are robust to other measures, such as entropy or isolation.

[6]For example, 21% of Muslims live in neighborhoods that are $>80\%$ Muslim, while 16% of SCs live in neighborhoods that are $>80\%$ SC.

[7]The historical record suggests that rural Indians have been highly endogamous, such that village settlement patterns observed today have been static for decades, if not centuries.

The magnitude of the disparities is large. For example, increasing the Muslim share of an urban neighborhood by 10 percentage points is associated with a 5 percentage point (or 10%) decrease in the probability that the neighborhood has a public secondary school. The local disparities are also large relative to those at higher levels of magnitude. The Muslim disadvantage in access to public services comes almost entirely from cross-neighborhood disparities. For SCs, allocations at higher levels of aggregation (across districts, for example), are in some cases progressive (as suggested by Banerjee and Somanathan (2007)), but become far more regressive once the cross-neighborhood distribution is considered. Urban primary schools are the only public facility where SCs have greater access to public facilities than non-SCs — Muslims are significantly disadvantaged on all public facilities that we could measure.

Importantly, these neighborhood-level disparities are in many cases not apparent in aggregate data. Thus, a policy maker observing school allocation at the district level would both be unaware of the failure of past policies to equalize access to public services, but would also fail to see the continued need for equalization policy.

We interpret these results through a stylized model, where government is hierarchical, and each level of government chooses the allocation of public services across its subsidiary hierarchical regions. This framework formalizes what is and is not measured by aggregate data analysis (with district-level data, for instance), and at which levels of government inequalities in service provision arise. It also highlights the relationship between disparities at different levels of aggregation and the government institutions operating at each level. The disparities that disadvantage minority groups are most prevalent at the most local and most informal levels of government — within towns and panchayats. These are also the levels of government which operate with the least scrutiny, and at the greatest distance from the level at which affirmative action policies are codified.

Finally, we examine the relationship between residential settlement patterns and outcomes for the next generation. Focusing on young adults, we find that young people growing up in marginalized group neighborhoods have worse educational outcomes. The disadvantages are much more severe in Muslim neighborhoods than in SC neighborhoods, a result parallel to the worse public service provision in those neighborhoods. These disparities in outcomes are apparent and strong even after controlling for parent education, and household and neighborhood consumption.

These results motivate two additional questions. Why do lower caste and Muslim neighborhoods have worse public services? And are the poor outcomes in these neighborhoods *caused* by living in these neighborhoods? Definitive answers to these questions are beyond the scope of this paper, but the evidence from India and around the world offers some suggestions.

The most obvious explanation for both of these phenomena is that suggested by the U.S. literature: that state and private discrimination drives minorities into segregated neighborhoods, which are then under-serviced with public facilities. (Cutler and Glaeser, 1997; Ananat, 2011; Aaronson et al., 2021). The chief alternative hypothesis is that the patterns we observe are driven by selection: individuals with poor potential outcomes choose to select into neighborhoods with low quality public facilities, perhaps because the cost of living is lower or because they choose to live with members of their own social groups. Selection across neighborhoods could also reflect a low valuation of public services by individuals who live in those neighborhoods.

There are undoubtedly forces that drive selection. A substantial share of individuals of all social groups express preferences for living around members of their own group (Center, 2021). It is possible that Muslims may value schools and health clinics with different characteristics than non-Muslims.

But there is also evidence that points against selection. First, a substantial share of India's schools and health clinics have been built recently and post-date at least the settlement patterns that we observe in villages. Second, the qualitative evidence suggests that Muslims would in fact prefer to have greater access to public services, but are unable to extract them from local politicians (Jaffrelot and Gayer, 2012). Indeed, lower castes' increased access to services in aggregate has happened exactly as lower caste groups were mobilizing politically and creating political parties to represent their interests, something that Muslims have never done successfully (Banerjee and Somanathan, 2007; Aneja and Ritadhi, 2022). Third, there is substantial evidence of discrimination in both the private and public sectors against both Muslims and SCs; one practice with U.S. echoes is that landlords and home sellers often refuse to work with Muslim buyers and renters (Banerjee and Knight, 1985; Sachar Committee Report, 2006; Thorat and Attewell, 2007; Madheswaran and Attewell, 2007; Thorat et al., 2015). Fourth, the negative relationship between neighborhood MG share and educational outcomes holds up even after controlling for average neighborhood consumption and for parent education, suggesting that poor child outcomes in MG neighborhoods have causes beyond the direct effect of (observably) poor people choosing to live in those neighborhoods.

Finally, we note the long historical record of high-status groups justifying disparities in public service provision with the argument that members of low-status groups do not want or merit those public services anyway (Blank et al., 2004). These arguments have been used as a justification for allowing those disparities to persist, and in many cases the arguments were decisively debunked decades later after the damage was done. Similarly, homophily among both U.S. Black and White Americans was frequently invoked as a justification for U.S. policies like red-lining, which entrenched or increased disparities across races. Given the poor historical record of arguments such as these, we might demand a greater standard of evidence for hypotheses and policies that allow cross-group disparities in public service provision to persist, as compared with policies that aim to mitigate those disparities. Nevertheless, further research providing more rigorous tests of these hypotheses would be highly valuable.

Systematic analysis of access to public services at the neighborhood level in developing countries has been elusive because of a paucity of census data with neighborhood identifiers, a data limitation that holds in most developing countries. While several of India's major sample surveys contain neighborhood identifiers, they are not powered to measure neighborhood characteristics like minority shares, or have enough coverage of neighborhoods *within* cities to measure urban segregation. Prior work includes a number of ward-level studies that use spatial units with population sizes of up to 30,000–200,000 people, 30 times the size of neighborhoods in our analysis.[8]

[8] Vithayathil and Singh (2012) use ward data from 2001 to show that residential segregation by caste is more prominent than by socioeconomic status in seven major cities. Singh et al. (2019) examine changes in

A series of recent studies has used enumeration block data similar to ours to document average patterns of segregation in a subset of Indian cities,[9] but we are aware of no prior work studying public service provision or individual outcomes at the neighborhood level in India, or any other major developing country. Even at the village level in India, there has been little prior analysis of disparities in villages with many Muslims, because data on village Muslim shares were unavailable, and we are aware of no prior work systematically studying access disparities *within* villages.

2 Context and Background

2.1 Scheduled Castes and Muslims in India

India's Scheduled Castes (SCs) are historically endogamous groups who occupy the lowest tiers of the caste system. They have experienced occupational and social segregation for thousands of years; social norms have effectively compelled them to take on low status occupations — like scavenging, emptying of toilets, or handling animal carcasses — with virtually no prospect of upward mobility. The practice of untouchability, now banned but still practiced in some form by many households, can take the form of segregation in schools, temples and markets, restrictions against entering the homes or even wearing sandals in the presence of higher caste groups, among others. These restrictions have been enforced with various social sanctions, including violence and murder (Girard, 2021). Since independence, the government of India has worked to mitigate the socioeconomic disadvantages of SCs through a range of programs and policies. SC status is often used as a marker of poverty for means-tested welfare programs, and there are reserved positions for SCs in higher education, politics, and in government. SCs still experience substantial socioeconomic disadvantages, but by many measures the gap between SCs and forward castes has shrunk somewhat over recent decades (Hnatkovska et al., 2012; Emran and Shilpi, 2015; Cassan, 2019; Asher et al., 2021a).

caste-based segregation from 2001 and 2011, again at the ward level, finding that residential segregation by caste has persisted or worsened in 60% of the cities in their sample. Neither of these studies examine religion, which is rarely available in Indian Census microdata.

[9] Bharathi et al. (2018) report enumeration block-level segregation based on SC status for five major cities. Bharathi et al. (2021) use similar scale data on caste and religion to characterize segregation in urban Karnataka. Susewind (2017) measures Muslim segregation using microgeographic polling booth data in eleven cities.

Muslims occupy a similar share of the population to Scheduled Castes (14% for Muslims vs. 17% for SCs). Like SCs, they on average have lower socioeconomic status than non-Muslim non-SCs. However, they experience fewer constitutional protections and have not been targeted by affirmative action policies. While SCs have been gaining ground on forward castes in socioeconomic terms, Muslims have if anything been losing ground, particularly in educational attainment, and have experienced significant losses in upward mobility in recent decades (Asher et al., 2021a). Post-independence India has been characterized by waves of anti-Muslim activism, sometimes resulting in riots, property destruction, and violence. Various social movements and political parties have mobilized around the idea of Hinduism as a key pillar of Indian identity, to the exclusion of Muslims (Jaffrelot, 2021). Our analysis uses data from 2011–13, and thus predates the rise of the current Modi regime (which has roots in these social movements), though the BJP (Modi's party) held power nationally in the early 2000s, and held power in many states before and during our sample period. Muslims have a higher share of members living in urban areas than any other major social group.

While SCs and Muslims represent the largest disadvantaged groups in India, there are several other social groups not separately considered by our analysis. Other Backward Castes (OBCs) occupy an intermediate place in the caste system between Forward Castes and SCs, comprising 40% of the population; IHDS 2011 reports that about half of Muslims are OBCs, though this share varies substantially across years and surveys. OBCs are not coded as such in any of the datasets that we use and their names are less distinctive, making it difficult to identify them (or their prevalence in any neighborhood) in our data. We also exclude Scheduled Tribes (STs) from our analysis; they are among the poorest social groups in India, but are concentrated in rural areas and have very small population shares in the vast majority of cities.[10] Given the focus of this paper, we use the terms "marginalized groups" or "minority groups" to describe SCs and Muslims, even though there are other groups in India that could also be classified as such.

[10] Only 4% of urban Indians report Scheduled Tribe status, compared with 15% who are SCs and 17% Muslims.

2.2 Minority Settlement Patterns in Rural and Urban India

Pre-independence cities in India were often characterized by neighborhoods with homogeneous occupational groups, often with mixed religion. In the absence of an effective municipal state, these neighborhoods were self-governing with respect to many public services, sometimes including even self-defense. Many had only a small number of entries, which made it possible to restrict access; this structure persists in many urban neighborhoods today, resulting in distinct boundaries between neighborhoods (Gist, 1957; , n.d.; Gould, 1965; Doshi, 1991; Chakrabarti et al., 2002).[11] The ethnographic literature suggests a secular trend of increasing segregation by religion rather than by occupation, as Hindu-Muslim violence has reduced Muslim feelings of safety in mixed neighborhoods. These newly concentrated Muslim neighborhoods can house individuals from many classes, often with income segregation existing within the neighborhoods at a smaller scale. Jaffrelot and Gayer (2012) describe this pattern of Muslim segregation in a series of monographs spanning many parts of the country. In many of the case studies, Muslims report difficulty getting attention from politicians or access to public services in their segregated neighborhoods.

The literature on villages also suggests a high degree of spatial separation between different classes and religions; individuals from lower status social groups often live in hamlets that are separated by a moderate walking distance from the village's primary agglomeration, where schools and health clinics typically appear (Beteille, 2012; Lanjouw et al., 2018).

While these patterns can be observed in many parts of the country, they are primarily documented in a qualitative literature (some of which is cited above), due to a general absence of data with neighborhood identifiers or of sufficient scale to characterize neighborhoods individually. There is a quantitative literature on unequal access to public services by caste *across* villages (Banerjee and Somanathan, 2007; Bailwal and Paul, 2021), in part because the decennial Population Census records the SC population share of every village, along with a series of public services. Nationwide data on village-level Muslim shares did not exist before this paper,

[11] These closed neighborhoods are described by different terms throughout the country: pols in Ahmedabad, mohallas in much of North India, paras in West Bengal, etc., often with names that reflect the occupational origins of the space. Muchipara, for instance, literally translates as "the neighborhood (para) of cobblers (muchi)."

nor was there data on either social group shares or public services at the neighborhood level *within* villages. To our knowledge, there has also been no large-sample study of public service variation *within* cities; an innovation of this paper is assembling urban neighborhood-level data on both public services and SC and Muslim shares.

2.3 Levels of government in India

India has a federal system of government with major powers divided between center, state, and local governments. The administrative apparatus is also decentralized, such that officials at different hierarchical levels have substantial autonomy.

There are 36 states and union territories (35 at the time of our sample), which have substantial administrative and legislative power. Public services are financed and allocated by both central and state government programs. Program implementation often lies in the hands of District Collectors, who are the top administrative officers of districts; there were 640 districts in our sample, though an additional 100 have been subdivided since then.

Local governance bodies are called panchayats in villages and municipalities in towns and cities. These bodies have elected representatives who can substantially influence the selection and allocation of public services within their administrative areas, but have little control over their overall budgets, most of which derive from grants from higher levels of government.

The most high-profile policies intended to close disparities between marginalized and non-marginalized groups are conceived and designed at the state and federal level, and often prescribe allocations of public services across aggregated levels of government. For instance, the District Primary Education Programme (Khanna, 2022) targeted funds for building schools to districts with below-median female literacy. The placement of new public facilities within districts, towns, and villages is rarely prescribed by these high-level policies; it is agreed upon through consultation with local elected leaders and bureaucrats. The extent to which policies target certain groups can therefore be different at different levels of aggregation; the less formal decision-making process of local bodies could either enhance or undermine the progressivity of policies designed at higher levels of government.

3 Neighborhood-level Data on Social Groups and Public Services

3.1 Identifying Neighborhoods

Studying neighborhood-level disparities requires data with granularity (to be able to identify neighborhoods) but also with scale (to be able to accurately measure neighborhood-level minority shares in many neighborhoods and many cities). Few of India's major sample surveys achieve this; they typically cover a small fraction of neighborhoods in any city, and too small a number of households in each sampling unit to measure minority shares or disparities in minority outcomes.

To bridge this gap, we combine a set of census data sources which use the internal survey block identifiers (enumeration blocks) that were created for the administration of India's 2011 Population Census. These "neighborhoods" consist of 100–200 households each, and describe a compact cluster of residences meant to be efficient for an enumerator to visit in a single session of work. In cities, they are typically city blocks or single buildings, while in rural areas their boundaries are typically defined by grouped clusters of residences. When villages consist of fewer than 100–200 households, an enumeration block is a single village.[12] Urban enumeration blocks are thus uniformly around 100–200 households, while rural blocks range from just a handful of households to the same upper limit around 200.[13] We exclude outlier neighborhoods which have fewer than 150 people (typically very small villages) or more than 1200 people.[14]

Note that "enumeration blocks" are not the same units as "census blocks" (sometimes just called "blocks"). There are 5000 census blocks in India with a mean population of about 200,000 each; these are a different level of aggregation used in some datasets and administrative programs. In this paper, we refer only to enumeration blocks.

While rural and urban enumeration blocks are similarly sized in terms of the number of

[12] About half of villages consist of a single enumeration block.

[13] Appendix Figure A.2 shows the distribution of block population in the sample. Results are very similar if we exclude rural villages that are too small to have a 100 household enumeration block, an analysis which results in similar distribution of urban and rural blocks.

[14] These make up less than 1% of the population and our results are unchanged if they are returned to the sample. Very large enumeration blocks are excluded because they are anomalous (suggesting potential data collection issues), and because segregation measures are scale-dependent (see below). However, there are so few of these neighborhoods, that the potential bias here would be small even if they were included.

households, the geography of urban and rural access to public facilities are quite different. Urban areas are dense, such that individuals may travel across many enumeration blocks for work or access to public services.

Rural areas are naturally more dispersed: neighboring enumeration blocks are separated by larger distances than in cities. Indian villages often contain distinct hamlets, which can be separated by a few hundred meters or more from the main settlement. It is anecdotally common to have a hamlet inhabited by lower castes or Muslims, which tends to be located at some distance from the village center. Because enumeration block boundaries are defined for the convenience of enumerators, multi-block villages with multiple hamlets will typically have enumeration block boundaries that keep hamlets self-contained within blocks. Villages tend to be much further apart from each other, with distance that varies across the country. The geocodes describing locations and polygons of enumeration blocks are not available to us — they are sold as hand-drawn maps at high cost by the Indian Census.

3.2 Public Facilities

The Population Census town and village directories report a wide range of public services, but these are only identified at the town/village level and map on to neighborhoods only for very small villages. To identify public facilities at the neighborhood level, we instead use the 2013 Economic Census (EC13). EC13 is a complete enumeration of non-farm establishments in the country, which includes schools, clinics, and hospitals, which are separately coded as private or public. EC13 records enumeration block identifiers, making it possible to identify public health clinics, primary schools, and secondary schools at the neighborhood level.[15] EC13 also records whether a firm owner is Muslim or SC. The employment share in SC or Muslim firms is highly correlated with the group share in each neighborhood.

We measure public service availability with binary measures that indicate whether an enumeration block contains a given type of public facility. Since individuals living in towns can

[15]The earlier rounds of the Economic Census (1990, 1998, 2005) record similar data, but with neighborhood identifiers (urban frame survey units) that do not match any census. It is thus not possible to study changes in neighborhood-level services over time.

more easily travel across many enumeration blocks, the consequence of having a public facility in your neighborhood may be smaller; nevertheless, disparities in access can still be interpreted as an outcome of the political economy process that determines where goods are allocated.

3.3 Demographic Data

Data on individuals comes from the Socioeconomic and Caste Census (SECC), a national asset census which recorded information on every household and individual in India (mostly in 2012) to determine eligibility for social programs. The SECC describes age, gender, education, occupation, and caste for every household member, as well as a short list of assets used to rapidly assess socioeconomic status. The SECC was made publicly available online in a combination of formats; we scraped and processed the data, with the approach described in detail in Asher and Novosad (2020). The urban data were not posted in their entirety; our sample covers 160 million urban residents, compared with the census urban population of 370 million. The set of towns and cities for which we have data are representative on the basis of population and urban amenities, and they span the entire country.

Consumption is not directly measured in the SECC, but we generate small area estimates of household per capita consumption on the basis of all of the household assets and characteristics on the SECC schedule, using the IHDS-II (2011–12) survey as our data source for consumption (Elbers et al., 2003). This process generates similar rural and urban consumption distributions to direct survey measures; see Asher and Novosad (2020) and Asher et al. (2021b) for more details. We aggregate per capita consumption to the enumeration block level as a proxy for neighborhood-level living standards; we also calculate these by social group.

The SECC surveyed individual caste and religion, but religion was not posted with the public data.[16] We therefore classify individuals as Muslims or non-Muslims using their first and last names, which were posted in the public data. Specifically, because of the distinctive naming patterns of Muslims, we can identify Muslim names with an out-of-sample accuracy

[16]Subcaste, or jati, was also recorded but not posted. The only caste identifier are broad indicators for Scheduled Caste or Scheduled Tribe status.

of 97%. We do this with an LSTM neural network, which classifies names on the basis of repeated letter sequences, using a religion-labeled dataset of 2 million applicants to the Indian Railways as a training sample. This approach has much higher accuracy than a fuzzy merge; the latter creates classification errors when small letter substitutions change a name identity, such as *Khan* (a stereotypical Muslim name) vs. *Khanna* (a Hindu name). The neural network implementation is described in detail in (Ash et al., 2022); a similar approach is taken by (Chaturvedi and Chaturvedi, 2020). We verified the classification accuracy on a withheld subset of the names in the railway data, as well as on a set of manually classified names in the SECC. Our classification also closely predicts the subdistrict-level population share of Muslims (Appendix Figure A.1). We pool Hindus with the 6% of Indians who are Jain, Christian, Sikh, or some other non-Hindu religion; the non-Hindu groups are small and we do not yet have an algorithm that can accurately classify them on the basis of names.

Finally, for comparison with the United States, we use data from the Diversities and Disparities project, which is based on the 2010 U.S. Census.

4 Methods

4.1 Measuring and Comparing Residential Segregation

To measure segregation of Muslims and SCs, we use the canonical dissimilarity index. The dissimilarity index ranges from zero to one and answers the question: what share of the minority group would need to change neighborhood for it to be evenly distributed within a city?[17] We calculate this index D_c for minority group G in city c as:

$$D_c = \frac{1}{2}\sum_{i=1}^{I}\left|\frac{G_i}{G_{\text{total}}} - \frac{G'_i}{G'_{\text{total}}}\right|, \tag{1}$$

where G_i is the number of members of group G in neighborhood i and G_{total} is the population of that group in the city. G' similarly refers to the non-group-G population.

[17] This conceptual understanding of dissimilarity treats housing units as mobile; moving a person to a neighborhood in this sense raises the population of the receiving neighborhood. It is nevertheless a useful way to conceptualize the measure.

When measuring Muslim segregation, we treat SCs as majority group members, and vice versa when measuring SC segregation. We take this approach because Muslim and SC segregation may have very different dynamics, causes, and consequences, suggesting that they should be considered separately. Tripartite segregation measures exist, but they do not describe the dynamics that we aim to explore here, as we are specifically interested in differences between SC and Muslim segregation. We conduct this pooling only when calculating segregation; when looking at access to public service and social group outcomes below, we always separate Muslims, SCs, and non-Muslim non-SCs.

For urban areas, we calculate the dissimilarity index for each city/town, defining enumeration blocks as neighborhoods. In rural areas, we calculate dissimilarity for each subdistrict, again with enumeration blocks as neighborhoods.[18]

Measures of segregation can change depending on the level of aggregation used to define neighborhoods. To take an extreme example, if we defined a "neighborhood" as a single household, we would calculate a dissimilarity index close to 1, given the very high rates of caste and religious endogamy. Our analysis defines neighborhoods at the enumeration block level (i.e. units of about 200 households or 1000 people), as these are the most accurate contiguous residential units that we can identify. This scale also fits our intuitive understanding of the set of households with which individuals will most often interact.

The scale-variance of the dissimilarity index means that a comparison with the United States — where census tracts have populations ranging from 1000 to 8000 and average around 4000 — would be biased toward finding greater segregation in India. Therefore, when we benchmark our segregation measures against the United States (and at no other place in the paper), we aggregate enumeration blocks based on their numeric identifiers to form neighborhoods of at least 4000 people.[19]

[18] A subdistrict consists of about 110 villages; there are about 5500 subdistricts in India. The rural measure thus captures a combination of segregation across villages and within villages.

[19] In the handful of cities where we have enumeration block maps or neighborhood names, we confirm that adjoining enumeration blocks are almost always adjoining in geography. Aggregating to 4000-person units based on block number inevitably adds a small amount of noise to the neighborhood definition, which is why we use the disaggregated neighborhoods for everything except the U.S. comparison. Note that the U.S. Census defines neighbor-

Our results are robust to aggregating neighborhoods to higher sizes, and the findings on segregation are robust to use of alternate segregation measures, like entropy or isolation.

4.2 Minority Shares and Neighborhood Public Services

Our second objective is to measure differences in access to public services between minority and non-minority neighborhoods. Given the different nature of governance at different hierarchical scales, we are interested in measuring how these differences vary by scale. We first present a stylized model of hierarchical government and public service allocation, and then describe how we estimate allocation disparities at different levels of aggregation.

4.2.1 A Toy Model of Hierarchical Government

We present a stylized description of how public services are allocated to neighborhoods in a system with multiple tiers of government. We describe the model with two levels of government, though we will estimate it at a higher level of granularity.

A national government (labeled I for India) chooses the allocation of spending on public services across districts. Each district therefore receives a fixed budget allocation; the district government (labeled D) chooses the allocation of that budget across neighborhoods (labeled N). All spending on goods takes place at the neighborhood level; each neighborhood spends all of its budget on a public service Y.

In reality, governments can influence the allocation across multiple levels of government; for instance, the federal government could target a grant to a specific city or place rules on how lower levels allocate funds. We abstract away from this possibility, and focus the model on the reality that each tier of government has a substantial degree of autonomy on the distribution of services across the level directly below it.

For simplicity, we assume that there are two districts (indexed 1 and 2), and each district has two neighborhoods indexed (1,1) and (1,2) in district 1, and (2,1) and (2,2) in district 2.

hoods according to existing informal boundaries, which are more likely to divide racial groups, overstating segregation relative to an approach studying random geographic units. Replicating this approach in India is not possible given the data available. As a result, the U.S. segregation measures will be biased upward relative to the Indian.

Each district and neighborhood is also characterized by a minority share of the population. Without loss of generality, we assume that district 1 has a higher minority share than district 2, and that the neighborhoods indexed 1 have a higher minority share than those indexed 2.

Let Y_1^D and Y_2^D be the budgets for districts 1 and 2 respectively. There is a budget constraint, such that total public spending $Y = Y_1^D + Y_2^D$. Let $Y_{d,1}^N$ and $Y_{d,2}^N$ be the budgets for neighborhoods 1 and 2 within district d, which must add up to the budget of district d: $Y_{d,1}^N + Y_{d,2}^N = Y_d^D$.

Each level of government is characterized by a political economy process which determines the allocation of spending across subsidiary levels, such that it allocates a share of total resources $(0.5+\alpha)$ to the subsidiary region with a higher minority share. The term α^I thus describes the pro-minority bias of national public spending (across districts), and α^D describes the pro-minority bias of district public spending (across neighborhoods within districts). When $\alpha > 0$, the budget allocation is pro-minority, and vice versa when $\alpha < 0$. We assume α^D is the same in both districts.

Formally, at the national level:

$$Y_1^D = (0.5+\alpha^I)Y$$

$$Y_2^D = (0.5-\alpha^I)Y,$$

And in each district d:

$$Y_{d,1}^N = (0.5+\alpha^D)Y_d^D$$

$$Y_{d,2}^N = (0.5-\alpha^D)Y_d^D.$$

We call α a political economy *function*, because it describes a complex and obscure political economy process that is a function of decisions by politicians, bureaucrats, firms, and citizens, which results in a scalar measure of disparities across locations. While we stylize α here as a choice made by some level of government, in practice, the minority share is also endogenous to individuals' location choices. When we estimate a proxy for α below, we thus interpret it not strictly as a choice of government, but as a statistic describing this entire political economy equilibrium. A negative value of α could therefore reflect discrimination by the state, or it could reflect historical inequalities that make minorities poorer and more likely to select into neighborhoods with worse public services. It describes the equilibrium inequality in public service allocation at one level of governance, but does not attribute it to a specific policy.

The overall allocation of public services to each neighborhood (i.e. the public services that individuals experience) is affected by both α^I and α^D:

$$
\begin{aligned}
Y_{1,1}^N &= (0.5+\alpha^I)(0.5+\alpha^D)Y \\
Y_{1,2}^N &= (0.5+\alpha^I)(0.5-\alpha^D)Y \\
Y_{2,1}^N &= (0.5-\alpha^I)(0.5+\alpha^D)Y \\
Y_{2,2}^N &= (0.5-\alpha^I)(0.5-\alpha^D)Y
\end{aligned}
$$

Extending the model to add additional levels of hierarchy is self-evident: the overall allocation of services to each neighborhood will depend on all of the α terms at higher levels of hierarchy. For cities, we will consider political economy functions at the levels of nation, state, district, and cities. For villages, we study allocation within the country, and then within states, districts, subdistricts, and villages.

This simple model draws attention to two features of the context that we study. First, the political economy processes may result in different allocations at different levels of government. The allocation of budgets across Indian states is highly formalized, scrutinized, and politically contested. The allocation of budgets across neighborhoods within cities or panchayats is less formal, may depend more on individual personalities, and decisions may not even be recorded.

Second, policy-makers and researchers may not observe allocations at all levels of the hierarchy. If neighborhood-level outcomes are not observed, then α^D is not observed, and the true level of disparity in public services is unknown.

4.2.2 Estimating the Political Economy Functions

We estimate the political economy functions at each level of aggregation indirectly. Specifically, we will estimate quantities α' which are proxies for the model α terms. The process is similar to a Blinder-Oaxaca decomposition, but where the covariates are hierarchical locations.

We describe our estimation in terms of urban location units (e.g. neighborhoods within cities) but the approach is identical for rural areas (neighborhoods within subdistricts). Each urban neighborhood N is embedded within a city C, a district D, a state S, and the country I. In terms of the model, the allocation of public facilities to that neighborhood depend on the set of political economy functions $\{\alpha^I, \alpha^S, \alpha^D, \alpha^C\}$.

Recall that the political economy function describes the effective choices made by a specific level of government, but affects the distribution of services at the level *below* it. So α^D is a function of district-level decisions, but affects the allocation of services across cities, not across districts.

We first estimate the relationship between the marginalized group share and the availability of a public service, at the neighborhood level, with city fixed effects:

$$SERVICE_{n,c} = \beta_c MG\_Share_{n,c} + \Omega_c + \boldsymbol{\nu} \boldsymbol{X}_{n,c} + \epsilon_{n,c}. \tag{2}$$

$SERVICE_{n,c}$ is a measure of the supply of public services in neighborhood n and city c, either an indicator for the presence of a public facility (a primary school, secondary school, or hospital), or the number of employees in a given type of public facility, the latter of which takes scale into account. Given that neighborhood size is relatively uniform, these two measures are very similar at the neighborhood level. The city fixed effect Ω_c controls for the allocation of public facilities *across* cities with more or fewer minorities, so that β_c describes the average distribution of public facilities *within* city. $\alpha'^{C} = \beta_c$ is thus a proxy for α^C.[20]

We include a control for neighborhood population, since it could be mechanically related to the supply of public services (though most neighborhoods are similarly populated). We also test robustness to a specification where SC and Muslim shares are included simultaneously, to test whether the group shares are independently associated with differential access to public services. We do not include other controls, because we do not want to control away any of the mechanisms which drive disparities between minority and majority neighborhoods. We cluster standard errors at the city level in the urban analysis, and at the subdistrict level in the rural analysis, to account for correlated outcomes within regions.

The city political economy function does not necessarily describe the overall disparity in public services experienced by minorities at the national level, because cities with higher or lower minority shares may have a higher or lower total supply of public services. However, it sheds light on the process of service allocation inside of cities.

To identify the political economy function at a higher level of aggregation, we decrease the granularity of the location fixed effect. Specifically, we can replace the city fixed effect with a district fixed effect as follows:

$$SERVICE_{n,c,d} = \beta_d MG\_Share_{n,c,d} + \Omega_d + \boldsymbol{\nu X}_{n,c,d} + \epsilon_{n,c,d}. \tag{3}$$

This estimation is still at the neighborhood level; the outcome is the measure of public facilities in neighborhood n in city c and district d. The additional subscript is added for clarity; the only difference from Equation 2 is that the city fixed effect Ω_c is replaced with a district fixed effect Ω_d.

In this equation, β_d describes the neighborhood-level relationship between minority share and access to public facilities, controlling for the allocation of public facilities across districts. The difference between β_c and β_d thus tells us how the distribution of public services across high- and low minority-share *cities* affects the ultimate distribution across neighborhoods. The

[20] We estimate a linear relationship for parsimony, but we will also present non-parametric graphs of the relationship, to visualize whether the relationship changes when the minority share is particularly high.

difference $\alpha'^D = \beta_d - \beta_c$ is thus a proxy for the district political economy function α^D.

Consider the following stylized example for clarification. Suppose the outcome is an indicator for the presence of a secondary school. If $\beta_d = 0.5$, it means that a 100% minority neighborhood is 50 percentage points more likely to have a secondary school than a 0% minority neighborhood, conditional on the number of secondary schools in each district. Suppose then that cities with more minorities have more schools — i.e. the district political economy function is pro-minority — so that controlling for city fixed effects brings the minority share coefficient down: $\beta_c = 0.3$. The change in coefficient $\beta_d - \beta_c = 0.2$ proxies the district political economy function α^D.

We repeat this process at two more levels of granularity. We replace Ω_d with a state fixed effect Ω_s; we obtain the state political economy function $\alpha'^S = \beta_s - \beta_d$. Finally, we estimate Equation 2 with no fixed effects at all, using the coefficient β_o; we obtain the federal political economy function $\alpha'^I = \beta_o - \beta_s$.

Note that β_o describes the average change in neighborhood public services as the minority share increases, controlling only for block population. This coefficient describes the disparity actually experienced by people living in marginalized group neighborhoods. Our estimation above allows this disparity to be decomposed into the the disparity at different geographic levels:

$$\beta_o = \alpha'^I + \alpha'^S + \alpha'^D + \alpha'^C$$

A useful feature of this decomposition is that it clarifies what information is lost by studying differences at aggregate levels. If we studied only the relationship between minority share and public service outcomes at the district level, we would calculate $\alpha'^I + \alpha'^S + \alpha'^D$, which may be a biased measure of β_o.

The estimations described here do not control for the economic status of people living in minority neighborhoods, and as such, the β coefficients cannot be directly interpreted as measures of discrimination or intentional targeting of public facilities based on social group. Nor do they isolate a causal effect of minority share on outcomes. For example, if minority groups are poor, and municipal governments undersupply public facilities to poor neighborhoods, then we would find $\beta_c < 0$ even if service provision is orthogonal to minority status, conditional on neighborhood income. However, if this was the case, it would remain true that minorities would experience worse access to public services—the outcome that we aim to measure.[21] Our null hypothesis is that the government allocates public facilities across neighborhoods equally, irrespective of neighborhood economic or social group status, in which case we would find $\beta_c = 0$.

We are interested in both the specific political economy functions at each level of aggregation, as well as the overall disparity represented by β_o. While β_o may be most directly relevant for minority welfare, it is important to understand the levels at which inequalities arise. If α^C is highly negative (i.e. discriminatory), average minority outcomes can still be improved by increasing α^S (e.g. through affirmative action programs operating across districts), but doing so will result in substantial misallocation. Distinguishing the different political economy functions is particularly relevant given the very different nature of the institutions governing the allocation of public services at different levels of aggregation, as described in Section 2. In particular, most policy research in India operates at the district level, as do many programs which determine the allocation of public services. High level policy-makers and researchers may not have access to systematic data describing the political economy processes at the local level, causing them to misunderstand the nature of inequalities.[22]

4.3 Minority Shares and Living Standards

Disparities in access to public services in minority neighborhoods are most concerning if they result in unequal outcomes for people living in those neighborhoods, as they can entrench inequality across groups. But if people in under-serviced neighborhoods can compensate by traveling to other neighborhoods for services or by consuming private services, then unequal allocation may be less harmful. The final part of our analysis therefore examines whether

[21] We do not necessarily get closer to causal identification by adding control variables for neighborhood average education or consumption, because these are outcomes which are plausibly caused by a shortage of public services.

[22] What is observed depends on the type of data being collected. If researchers observe only district school allocations and district minority shares, then they cannot see the outcomes of any political processes that allocate goods below the district level. If they observe individual-level data on whether people have access to public services, then they effectively see the sum of within-city, cross-city, and cross-district allocation, but they cannot see at which level the disparity originates.

individuals experience worse outcomes in minority neighborhoods.

In this section, we investigate whether individual outcomes are worse in minority neighborhoods. We use the following estimation equation to examine the relationship between neighborhood minority share and the young generation's educational outcomes:

$$ED_{i,n,c} = \beta MG\_Share_{i,n,c} + \Omega_c + \boldsymbol{\nu} \boldsymbol{X}_{i,n,c} + \epsilon_{i,n,c}. \tag{4}$$

The outcome is the education level of people aged 17–18, measured as an indicator for completion of a given schooling level. We control for parent education, household consumption, and average neighborhood consumption.

The analysis is at best suggestive of the causal relationship between neighborhood and child outcomes, and has a number of limitations. First, we do not know individuals' place of birth. We focus on young people, and test robustness by looking at the middle school completion of 16–17 year olds (who are very likely to still be at home), but we do not have the data to exclude recent migrants from the sample. More importantly, there may be unobserved characteristics of households which cause them to have worse outcomes and also choose to live in neighborhoods with other minorities. Finally, we may be over-controlling, in that average neighborhood consumption could be affected by the poor delivery of services to minority neighborhoods; this would bias down β relative to the causal estimator.

Nevertheless, these results can be useful even if interpreted entirely descriptively; they tell us the extent to which minority neighborhoods are places where children are getting appropriate educations. They can also motivate research into identification strategies which can better identify a causal estimator in this context.

5 Results

5.1 Segregation in Indian Villages and Cities

Table 1 presents demographic data at two levels of aggregation. Columns 1 and 2 show the characteristics of average rural and urban neighborhoods (i.e. enumeration blocks) respectively.

Both rural and urban blocks have about 500 people each; there are about 1.1 million rural blocks and 420,000 urban blocks in the sample. The difference in sample size reflects both India's low urbanization rate (31% in 2011), and the fact that our urban data only covers 80% of urban people in India, as noted in Section 3. Scheduled Castes are more concentrated in rural areas, while Muslims are more urban.

Most of the variation in where these groups live is highly local. Only 20% of the variation in the urban Muslim share is cross-district, and 70% is within-town. For urban SCs, these figures are 8% and 12% respectively. That variation in group share is so local is already a sign that differences in public goods allocations at high levels of aggregation will have muted effects on neighborhood access. In rural areas, aggregate regions are slightly more predictive of where SCs and Muslims live. For Muslims, 65% of variation is within subdistrict, and 29% is within-village; for Muslims, these figures are 85% and 55%.

Differences in segregation across the two groups are difficult to capture in a single dimension. Figure 1 shows the distribution of the dissimilarity index across cities and across rural subdistricts. In both charts, the distribution of dissimilarity for SCs is systematically higher than that for Muslims, confirm the statistic reported in Table 1. The urban and rural distributions are highly similar, but note that these indices measure clustering on different geographic scales.[23] Panel B overlays the distribution of segregation of Black people in the U.S.; the distribution is highly similar to that of Muslims in urban India.

Figure 2 shows an alternate view of residential segregation, by showing the distribution of each minority population across neighborhoods with different minority shares. In both urban and rural areas, the Muslim distribution is highly bimodal, with many Muslims living in highly segregated neighborhoods, but many Muslims also living in highly integrated neighborhoods. Scheduled Castes, meanwhile, mostly live in moderately segregated neighborhoods, but very few are as clustered as the most segregated Muslims. Paradoxically, SCs are more integrated on average, but a greater share of Muslims live in highly segregated neighborhoods. In cities, the

[23]The urban measure describes segregation across neighborhoods, which are very close to one another. The rural measure describes segregation across both neighborhoods and villages; the latter are much more dispersed.

median Muslim lives in a neighborhood that is 46% Muslim. In rural areas, this is 35%. For SCs, these numbers are 37% and 45%, almost exactly the reverse. 21% of urban Muslims live in neighborhoods that are >80% Muslim, while only 16% of SCs live in neighorhoods that are >80% SC.

Figure 3 shows a binned scatterplot comparing measures of rural and urban segregation across districts; they are highly correlated for both SCs and Muslims ($\rho = 0.57$ in both cases). Cities appear to be replicating the segregation patterns of their hinterlands.

Figure 4 shows maps of SC and Muslim segregation across the country. While there are pockets of high and low segregation, they do not follow obvious geographic patterns; the north is no less segregated than the south. Table 2 shows summary statistics for public services at the neighborhood level (Columns 1 and 2), and at the town (urban) and subdistrict (rural) level in Columns 3 and 4.

5.2 Allocation of Public Facilities across Neighborhoods

Having established that segregation of Muslims and Scheduled Castes is substantial in both urban and rural areas, we now examine how the supply of public services varies across neighborhoods with and without large populations from these marginalized groups.

We begin by non-parametrically examining the most local political economy functions, those of the city and the rural subdistrict. Figure 5 shows a binned scatterplot of the neighborhood-level relationship between the supply of secondary schools (an indicator for the presence of a neighborhood school) and the neighborhood marginalized group share, for each group, in both urban and rural areas. The urban series is residualized on city fixed effects and thus describes the city political economy function—it describes how schools are distributed across neighborhoods, conditional on the total school supply in a city. Secondary school availability falls monotonically with the neighborhood Muslim share; raising the Muslim share of a neighborhood by 50 percentage points is associated with a 28% lower likelihood of the neighborhood having a public secondary school (approximately a 0.6 percentage point decline on a mean of 2.4%). Neighborhoods with a ¿50% Muslim share stand out for being underprovisioned; there are not so many of these neighborhoods in India, but as shown in Figure 2, a large share of Muslims live in them. Rural locations look broadly similar, with the most Muslim neighborhoods having substantially fewer schools.

The relationship between Scheduled Caste share and secondary school access appears to be non-monotonic; at low levels of SC shares, it is flat or rising, but above a 20% SC share, secondary school presence falls precipitously, such that 50% SC neighborhoods have similar school availability to 50% Muslim neighborhoods. The rural plots again look broadly similar to the urban.[24]

We summarize this nonparametric relationship between local minority share and local public facilities with the linear estimator given by Equation 2; this is α^C. Table 3 shows, for urban areas, these linear regression estimates for each kind of public facility that we can measure: primary schools, secondary schools, and hospitals, always with city fixed effects. The dependent variable in odd numbered columns is an indicator for facility availability, while in even-numbered columns it is log employment in public facilities of the given type, a measure of scale.

Table 4 shows the same result for rural areas. SC and Muslim shares are included simultaneously to ensure that the allocation of facilities to one group's neighborhoods does not drive our estimate for the other group. The tables confirm that at the most local geography, SC and Muslim neighborhoods are systematically allocated fewer public services; with the exception of urban primary schools in SC neighborhoods, the point estimates are all negative, substantial, highly significant. Moreover, for every public service, in every location, Muslim neighborhoods are less well served than SC neighborhoods.

These results describe only the local political economy of public services; they do not describe the service availability faced by Muslims and SCs, which depends also on the higher-level political economy functions. The full set of political economy functions is summarized by Figure 6 for Muslims and Figure 7 for Scheduled Castes. We take some time with these figures as they describe a central result of this paper.

We walk through Panel A of Figure 6, which shows the political economy functions for the supply of primary schools to urban neighborhoods as a function of the Muslim share. The outcome variable is the number of primary schools per 100,000 people; the sample mean of this

[24]Rural school shares are higher on average because rural areas are characterized by a greater number of smaller schools, reflecting the greater distance between neighborhoods.

variable is 15. The rightmost gray box tells us that a 100% Muslim neighborhood is estimated to have 2.2 fewer primary schools per 100,000 people than a 0% Muslim neighborhood.[25] This estimate is β_o from Equation 2, estimated without fixed effects. This is the total disparity experienced by people living in such a neighborhood.

This gap of 2.2 fewer primary schools can be decomposed into the different levels of aggregation at which it appears. The leftmost estimate shows $\alpha^I = -0.5$; this means that 0.5 points out of the 2.2 point disparity arise from states with more Muslims having fewer primary schools. The second estimate from left gives us $\alpha^S = +1.1$: conditional on the number of primary schools in a state, districts with more Muslims on average have more primary schools. The next two bars respectively give us α^D and α^C, which both have regressive allocations. The sum of all the α coefficients gives us the final estimate of -2.2, which is driven almost entirely by the allocation of primary schools across urban neighborhoods within towns.

We highlight several features of the results in Figure 6. First, the city political economy functions are systematically negative — these are just the graphical representations of Tables 3 and 4. Second, the magnitude of the city political economy functions dwarfs the magnitude of the political economy function at every other level of aggregation. For four out of the six public facilities, the overall disparity in Muslim neighborhoods is driven primarily by differences *within* towns and villages. The lowest and most informal level of governance is where Muslim neighborhoods are most left out. Third, without neighborhood level data, we would detect no disadvantage in access to public facilities for Muslims.

Figure 7 shows the same results for SC neighborhoods. The patterns are distinct from those observed for Muslims. A clear pattern emerges for secondary schools and hospitals, in both rural and urban areas. The allocation of these services is neutral or progressive across states, districts, towns, and villages. But within towns and villages, the distribution of schools and hospitals across neighborhoods is highly regressive, undoing almost all of the progressivity at higher levels of government. Urban primary schools have progressive allocations at all levels

[25]The sample means for the other variables are in the figure note.

of aggregation, while rural primary schools are negative at all levels. Across all types of public facilities, both the overall allocation and the cross-neighborhood allocation are substantially more regressive as a function of the Muslim share than they are of the SC share.

One explanation for these patterns is that affirmative action policies for Scheduled Castes primarily affect the distribution of public services across higher units of aggregation, like states and districts. These policies may have less traction for the less formal processes of service allocation across neighborhoods; if these processes remain biased against marginalized groups, the cross-neighborhood allocation will undo some of the progressive allocation at higher levels of government. For Muslims, we see the same disadvantages at the cross-neighborhood level, but there has been no systematic policy of affirmative action that would drive progressive allocation of public services to Muslim regions at higher levels of aggregation.

5.3 Minority Economic Outcomes in Segregated Neighborhoods

In this section, we examine the relationship between young people's educational outcomes and the minority share of the neighborhood where they live, using Equation 4. Table 5 shows the results. The coefficients of interest are the Scheduled Caste and the Muslim shares, estimated simultaneously. Results are shown separately for boys (Columns 1–3) and girls (Columns 4–6). In each set of three columns, we first show an estimate with no controls (Columns 1 and 4), then with controls for parent education (Columns 2 and 5), and finally with controls for parent education, parent consumption, and neighborhood consumption (Columns 3 and 6). All columns include city (or subdistrict, for rural areas) fixed effects; these results are strictly across neighborhoods within cities and subdistricts.

We highlight two results from this table. First, across all comparisons, young people's educational outcomes are worse in Muslim neighborhoods than they are in SC neighborhoods, and considerably so. In rural areas, the Scheduled Caste share is not even statistically correlated with the next generation's educational outcomes; in urban areas, with all controls in place, the coefficient on Muslim share is almost ten times larger than that on the SC share. An individual in a Muslim enclave is predicted to obtain about 0.1 fewer years of education than an individual in an integrated neighborhood.

Second, the coefficients remain strongly statistically significant and negative (though not for SCs in rural areas) after household and neighborhood controls are added. Observable family characteristics do not explain the disadvantages faced by young people growing up in Muslim neighborhoods in particular. Of course, there may be unobserved characteristics of families that can explain some of the remaining variation. However, these coefficients could also be downward-biased, if some of the control variables are in fact downstream of the neighborhood group share. For instance, if consumption was low, say, because individuals have poor security of tenure in Muslim neighborhoods, then it would be a bad control which would cause us to estimate less than the causal effect of the neighborhood on the child's educational outcome.

Finally, we can estimate these same regressions on samples restricted to members of given social groups. The resulting coefficients are summarized in Figure 8, with the full set of controls above. The results show that the findings are broadly similar across all groups; all children in Muslim-heavy neighborhoods have worse educational outcomes, not only Muslims.

6 Conclusion

This paper describes a national-scale analysis of socio-economic outcomes and access to public services in India's urban and rural neighborhoods. Analysis of this kind has previously been impossible on a large scale due to the absence of sufficient neighborhood-level data to characterize minority and non-minority neighborhoods.

India's growing cities are highly segregated. They are only marginally less segregated than rural areas, where neighborhood structure is strongly conditioned by centuries of occupation- and status-based division via the caste system. The religious and caste identity of the people who live in a given urban neighborhood are strongly predictive of both access to public services and of socioeconomic status in those neighborhoods. Both Muslims and Scheduled Castes are highly segregated, but Muslims experience much worse access to public services as a function of that segregation. India's rapidly growing cities, famous as engines of upward mobility, to a large degree have replicated the caste and religious structure of its villages.

Our research so far does not identify the causes of these disparities. However, discriminatory provision of public facilities to minority neighborhoods has a storied history in many countries, including in India. It is perhaps unsurprising to find it also existing at the neighborhood level, but the disproportionate disadvantage of Muslim neighborhoods, even relative to SC neighborhoods, is striking.

A limitation of our work is that it is based on cross-sectional data collected in 2012–13; it is challenge to document the time path of disparity. The historical literature suggests that Scheduled Castes have been isolated at the neighborhood level for generations, but Muslim isolation has been exacerbated by Hindu-Muslim violence in the post-colonial era. Data from historical censuses could potentially shed light on changes in residential segregation over time.

Consistent with the similarity we found between rural and urban areas, the limited evidence that we have suggests that India's residential patterns are not changing very rapidly. Younger cities in our sample are less segregated than older cities, but the pace of change is glacial; at this rate, India's Muslims and Scheduled Castes will not be integrated in cities for another 500 years.

That living standards are so much lower in SC and Muslim enclaves suggests that, as elsewhere, spatial concentration of marginalized groups may limit their economic opportunities. Modern India has never had the government regulations, such as redlining, that contributed to racial segregation in the United States — there are thus fewer overtly harmful policies to remove. However, housing discrimination in India's cities is widely documented and has even been explicitly tolerated by the judiciary, echoing patterns from a too recent era in the U.S..

The historic tolerance for residential segregation and unequal access to public services has had disastrous consequences for the United States; it has prevented generations of individuals from access to opportunity, and is a central fracture in a highly polarized political system. At an earlier stage of development and with cities still rapidly growing, India has the opportunity to make a different set of choices. By highlighting segregation in India and documenting the concomitant disparities in access to public services, we hope to draw attention to the critical choices ahead.

References

, in "in."

Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder, "The effects of the 1930s HOLC" redlining" maps," *American Economic Journal: Economic Policy*, 2021, *13* (4), 355–92.

Alesina, Alberto, William Easterly, and Janina Matuszeski, "Artificial states," *Journal of the European Economic Association*, 2011, *9* (2), 246–277.

Ananat, Elizabeth Oltmans, "The Wrong Side(s) of the Tracks: The Causal Effects of Racial Segregation on Urban Poverty and Inequality," *American Economic Journal: Applied Economics*, 2011, *3* (2), 34–66.

_ **and Ebonya Washington**, "Segregation and Black political efficacy," *Journal of Public Economics*, 2009, *93* (5-6), 807–822.

Aneja, Abhay and S.K. Ritadhi, "Can political parties improve minority welfare? Evidence from India's "Silent Revolution"," *Journal of Development Economics*, 2022, p. 102931.

Ash, Elliott, Sam Asher, Aditi Bhowmick, Sandeep Bhupatiraju, Daniel L Chen, Tanaya Devi, Christoph Goessmann, Paul Novosad, and Bilal Siddiqi, "In-group bias in the Indian judiciary: Evidence from 5 million criminal cases," 2022. Working paper.

Asher, Sam and Paul Novosad, "Rural roads and local economic development," *American Economic Review*, 2020, *110* (3), 797–823.

_ **,** _ **, and Charlie Rafkin**, "Intergenerational Mobility in India: New Methods and Estimates Across Time, Space, and Communities," 2021. Working Paper.

_ **, Tobias Lunt, Ryu Matsuura, and Paul Novosad**, "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*, 2021, *35* (4), 845–871.

Bailwal, Neha and Sourabh Paul, "Caste Discrimination in Provision of Public Schools in Rural India," *The Journal of Development Studies*, 2021, *57* (11), 1830–1851.

Banerjee, Abhijit and Rohini Somanathan, "The political economy of public goods: Some evidence from India," *Journal of development Economics*, 2007, *82* (2), 287–314.

Banerjee, Biswajit and John B Knight, "Caste discrimination in the Indian urban labour market," *Journal of development Economics*, 1985, *17* (3), 277–307.

Basant, Rakesh and Abusaleh Shariff, *Handbook of Muslims in India: Empirical and policy perspectives*, Oxford University Press, 2010.

Beteille, Andre, *Caste, class and power: Changing patterns of stratification in a Tanjore village*, Oxford University Press, 2012.

Bharathi, Naveen, Deepak Malghan, and Andaleeb Rahman, "A permanent cordon sanitaire: intra-village spatial segregation and social distance in India," *Contemporary South Asia*, 2021, *29* (2), 212–219.

_ , Deepak V Malghan, and Andaleeb Rahman, "Isolated by caste: Neighbourhood-scale residential segregation in Indian metros," *IIM Bangalore Research Paper*, 2018, (572).

Blank, Rebecca M, Marilyn Dabady, Constance F Citro, and Rebecca M Blank, *Measuring racial discrimination*, National Academies Press Washington, DC, 2004.

Boustan, Leah Platt, "Racial Residential Segregation in American Cities," *NBER Working Paper*, 2013, (No. w19045).

Cassan, Guilhem, "Affirmative action, education and gender: Evidence from India," *Journal of Development Economics*, 2019, *136*, 51–70.

Center, Pew Research, "Religion in India: Tolerance and Segregation," 2021.

Chakrabarti, Vibhuti, Vibhuti Sachdev, Giles Henry Rupert Tillotson, Giles Sachdev, and Vibhuti Tillotson, *Building Jaipur: the making of an Indian city*, Reaktion Books, 2002.

Chaturvedi, Rochana and Sugat Chaturvedi, "It's All in the Name: A Character Based Approach To Infer Religion," 2020. arXiv preprint arXiv:2010.14479.

Chetty, Raj and Nathaniel Hendren, "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects," *The Quarterly Journal of Economics*, 2018, *133* (3), 1107–1162.

_ , _ , and Lawrence F Katz, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 2016, *106* (4), 855–902.

Cutler, David M. and Edward L. Glaeser, "Are Ghettos Good or Bad?," *The Quarterly Journal of Economics*, 1997, *112* (3), 827–872.

Cutler, David M, Edward L Glaeser, and Jacob L Vigdor, "When are ghettos bad? Lessons from immigrant segregation in the United States," *Journal of Urban Economics*, 2008, *63* (3), 759–774.

Doshi, Harish, "Traditional neighbourhood in modern Ahmedabad: the Pol," *A Reader in Urban Sociology. New Delhi: Orient Longman*, 1991, pp. 179–208.

Elbers, Chris, Jean Lanjouw, and Peter Lanjouw, "Micro-level Estimation of Poverty and Inequality," *Econometrica*, 2003, *71* (1), 355–364.

Emran, M.S. and Forhad Shilpi, "Gender, Geography, and Generations: Intergenerational Educational Mobility in Post-Reform India," *World Development*, 2015, *72*, 362–380.

Girard, Victoire, "Stabbed in the back? Mandated political representation and murders," *Social Choice and Welfare*, 2021, *56* (4), 595–634.

Gist, Noel, "The Ecology of Bangalore, India: An East-West Comparison," *Social Forces*, 1957, *35* (4), 356–365.

Gould, Harold A, "Lucknow Rickshawallas: The Social Organization of an Occupational Category," *International Journal of Comparative Sociology*, 1965, *6* (1).

Gulzar, Saad, Nicholas Haas, and Benjamin Pasquale, "Does Political Affirmative Action Work, and for Whom? Theory and Evidence on India's Scheduled Areas," *American Political Science Review*, 2020, *114* (4), 1230–1246.

Hnatkovska, Viktoria, Amartya Lahiri, and Sourabh B. Paul, "Castes and Labor Mobility," *American Economic Journal: Applied Economics*, 2012, *4* (2).

Jaffrelot, Christophe, *Modi's India: Hindu Nationalism and the Rise of Ethnic Democracy*, Princeton University Press, 2021.

_ **and Laurent Gayer**, *Muslims in Indian cities: Trajectories of marginalisation*, Columbia University Press, 2012.

Khanna, Gaurav, "Large-scale Education Reform in General Equilibrium: Regression Discontinuity Evidence from India," 2022. Working paper.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Lanjouw, Peter, Nicholas Stern et al., *How lives change: Palanpur, India, and development economics*, Oxford University Press, 2018.

Madheswaran, Subramaniam and Paul Attewell, "Caste discrimination in the Indian urban labour market: Evidence from the National Sample Survey," *Economic and political Weekly*, 2007, pp. 4146–4153.

Sachar Committee Report, "Social, Economic and Educational Status of the Muslim Community of India," Technical Report, Government of India 2006.

Singh, Gayatri, Trina Vithayathil, and Kanhu Charan Pradhan, "Recasting inequality: residential segregation by caste over time in urban India," *Environment and urbanization*, 2019, *31* (2), 615–634.

Susewind, Raphael, "Muslims in Indian cities: Degrees of segregation and the elusive ghetto," *Environment and Planning A*, 2017, *49* (6), 1286–1307.

Thorat, Sukhadeo and Paul Attewell, "The legacy of social exclusion: A correspondence study of job discrimination in India," *Economic and political weekly*, 2007, pp. 4141–4145.

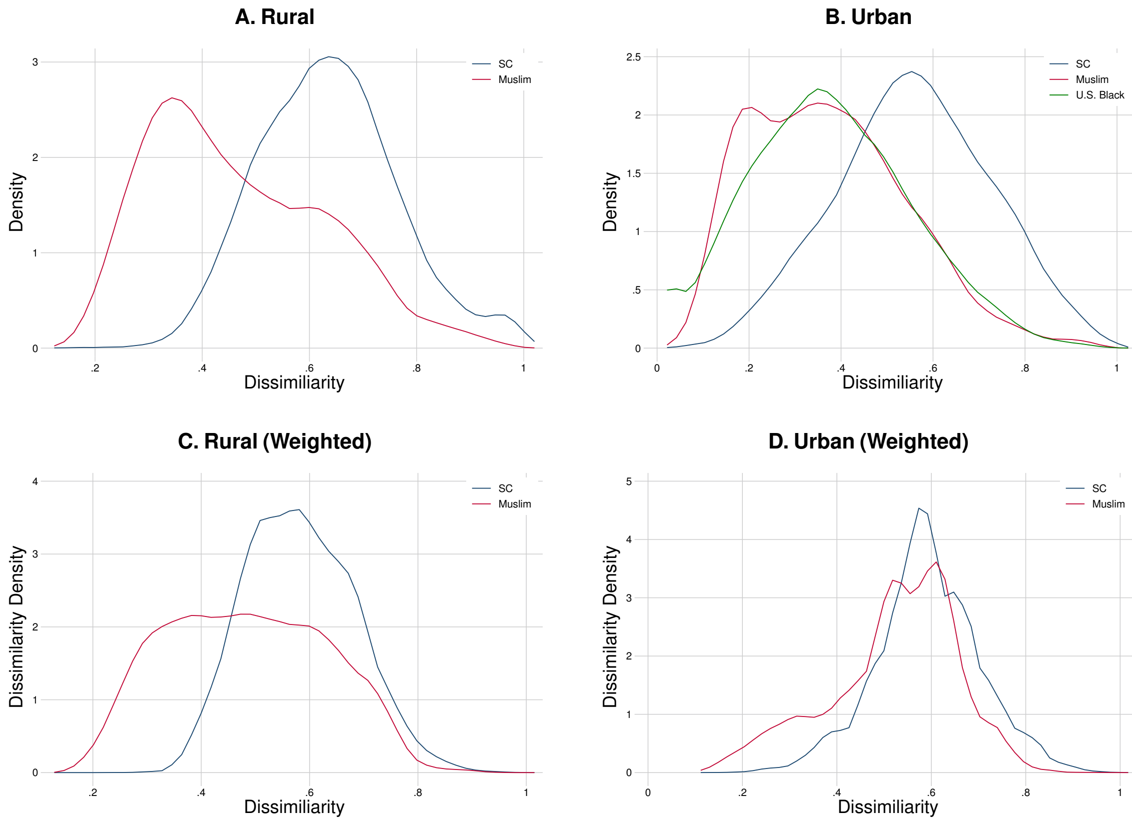
_ , Anuradha Banerjee, Vinod K Mishra, and Firdaus Rizvi, "Urban rental housing market: Caste and religion matters in access," *Economic and Political Weekly*, 2015, pp. 47–53.

Vithayathil, Trina and Gayatri Singh, "Spaces of discrimination: residential segregation in Indian cities," *Economic and Political Weekly*, 2012, pp. 60–66.

Figures

Figure 1

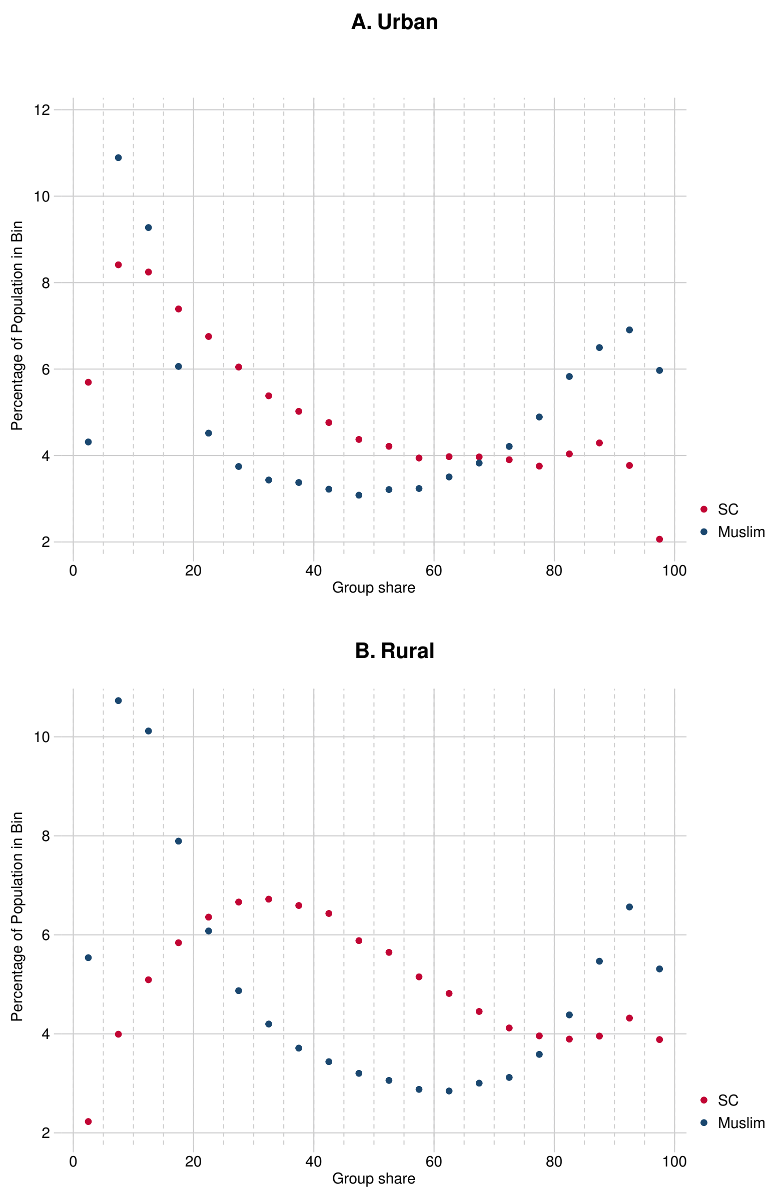
Extent of Segregation



Notes: This figure shows density plots of SC and Muslim dissimilarity. Panels A and C show the distribution of dissimilarity indices across rural subdistricts; Panels B and D show it across towns and cities. Panel B includes urban US Black dissimilarity. Panels C and D weight the density plots by city and subdistrict population respectively.

Figure 2

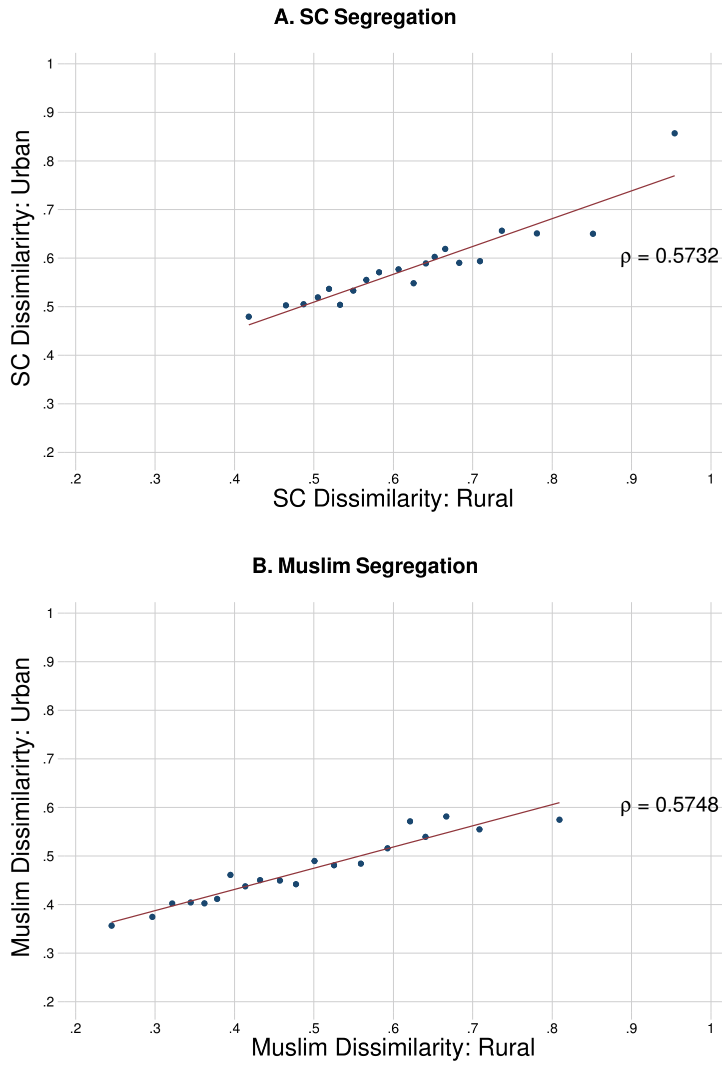
Population Distribution as a Function of Minority Share



Notes: The figure shows the distribution of Scheduled Caste and Muslim Population shares across their own neighborhood group share. For instance, the rightmost blue point in Panel A reveals that 6% (Y-axis) of Muslims live in neighborhoods where the Muslim share is between 95 and 100%.

Figure 3

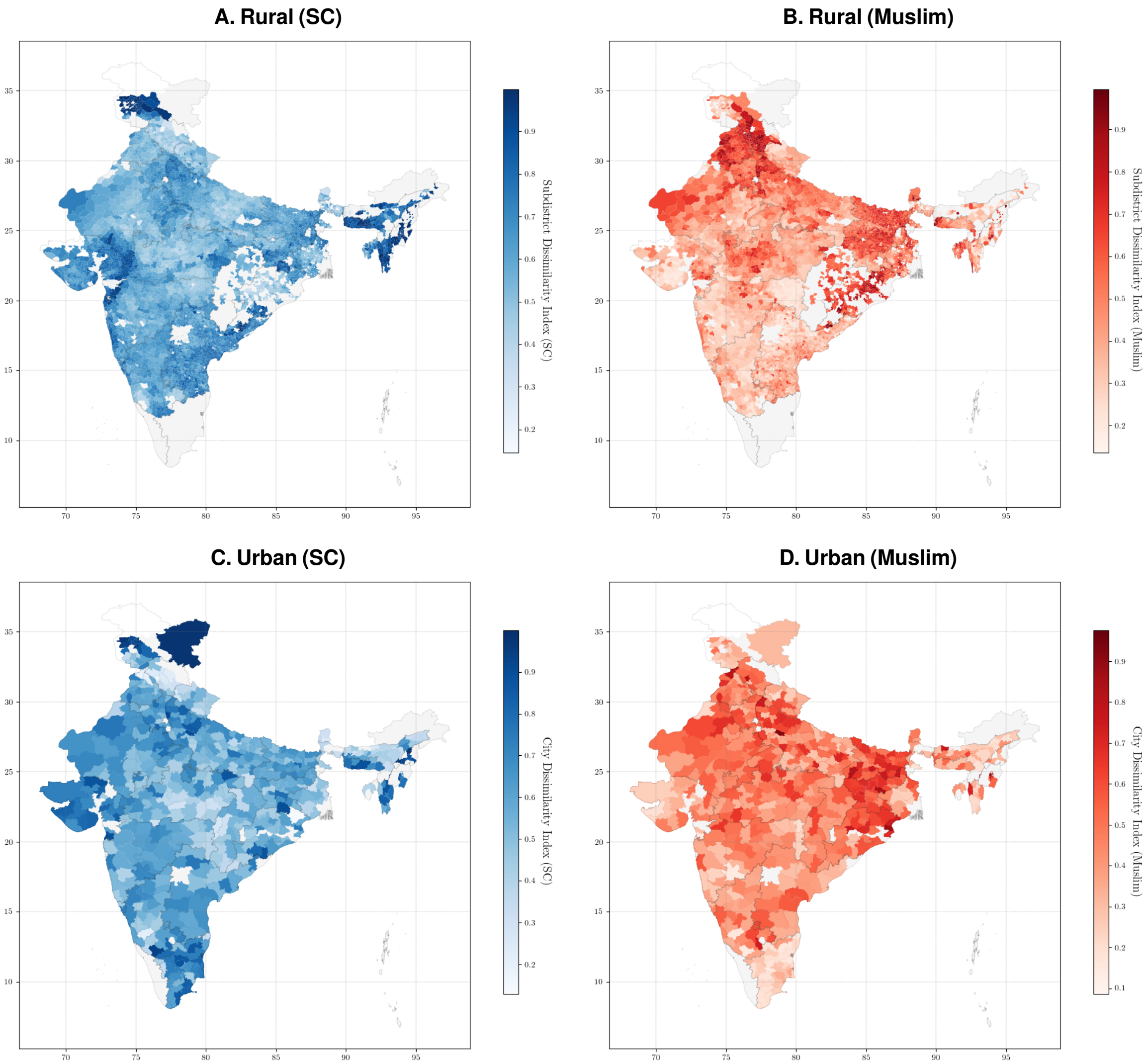
Urban vs Rural Segregation: District-level Comparisons



Notes: This figure shows a binscatter representing the district-level correlation between rural and urban segregation. It uses our standard dissimilarity measure of segregation at the town/subdistrict-level, but aggregated up to the district for comparison across rural and urban areas. Group shares and resultant dissimilarity indices come from the Socioeconomic and Caste Census.

Figure 4

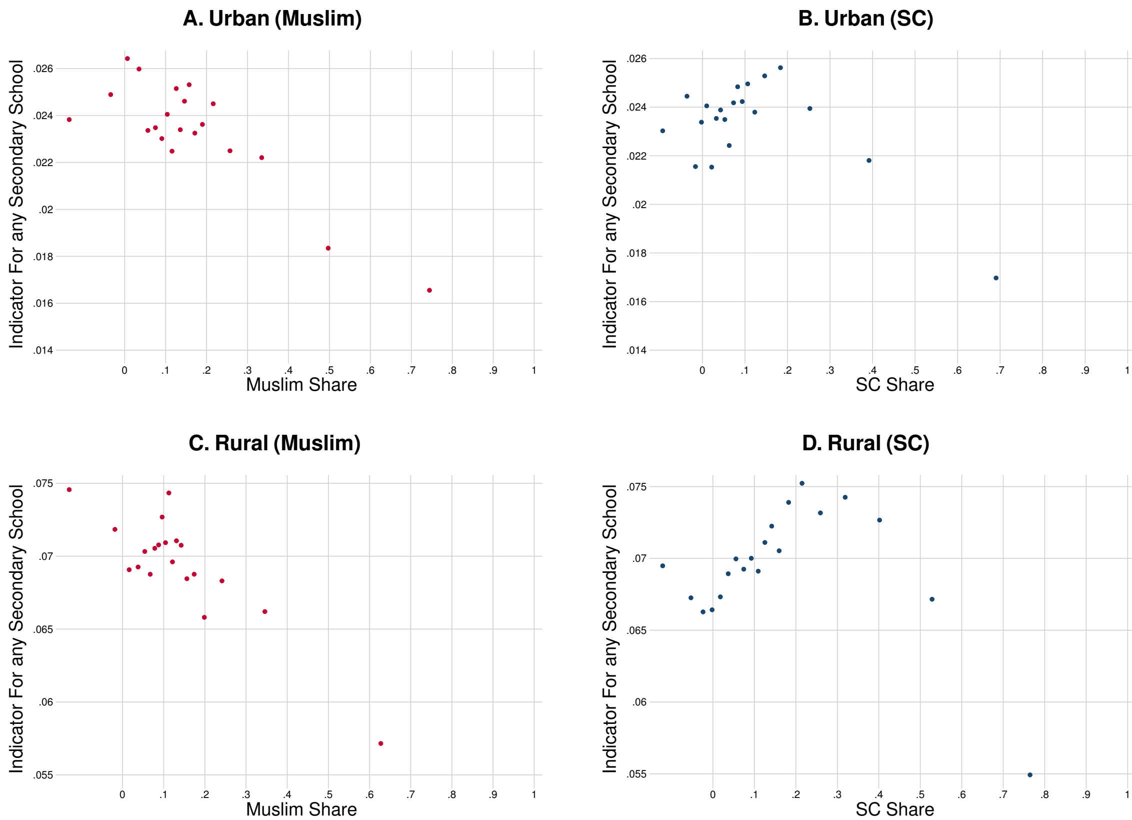
Segregation Maps



Notes: These maps show segregation across India. The top row presents heatmaps represent rural subdistricts for SCs (left) and Muslims (right). The bottom row does the same, but presents urban segregation aggregated to the district-level for visibility.

Figure 5

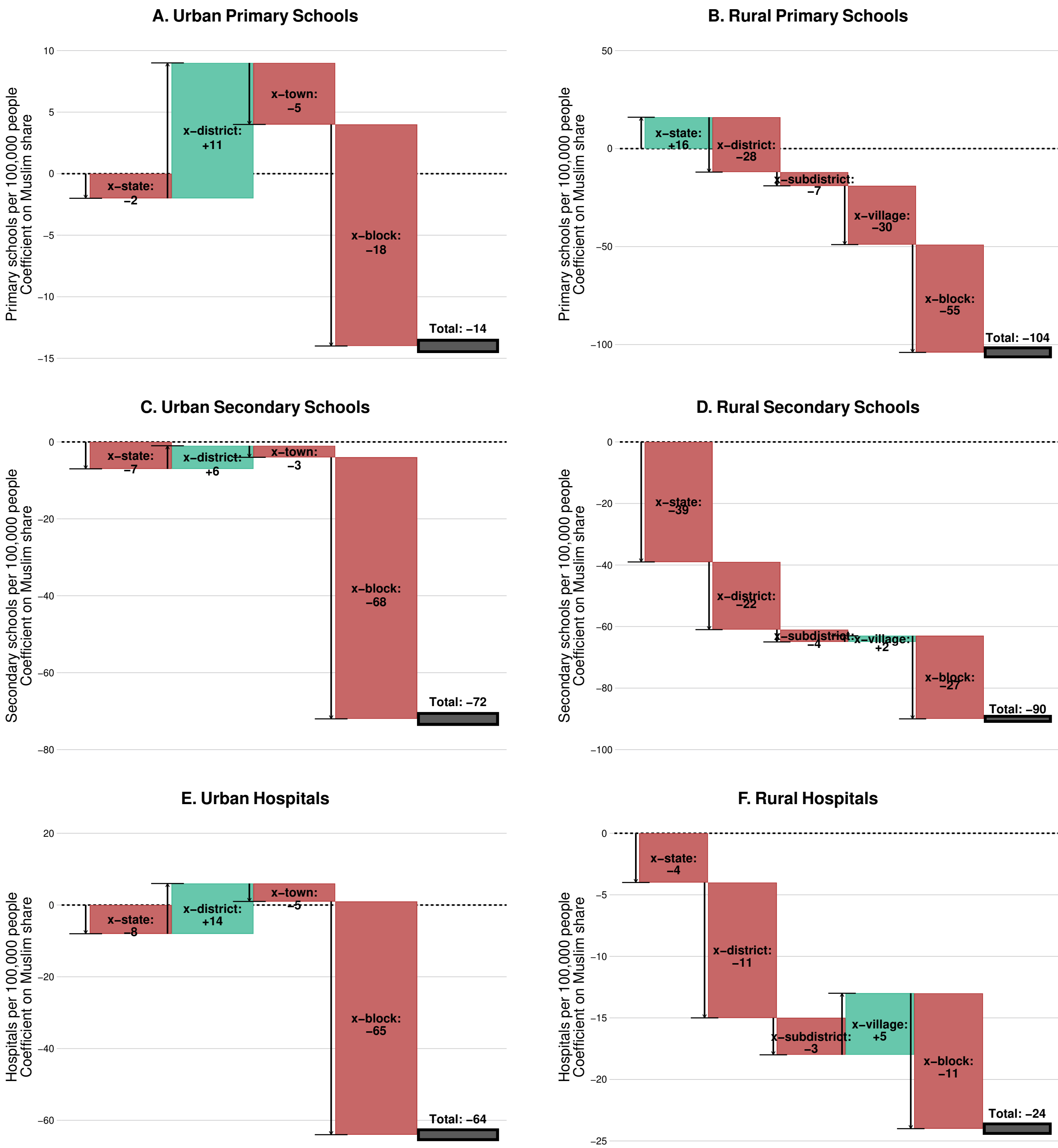
The City and Subdistrict Political Economy Functions:
Secondary Education Facilities vs Marginalized Group Shares



Notes: This figure shows binscatter plots of the percentage of neighborhoods that have a secondary school at a given level of SC/Muslim share. The top panels show this for rural areas and the bottom panels show these for urban areas. The panels on the top and bottom left show the binscatters of the percentage of neighborhoods with secondary schools with the neighborhood SC share, and the panels on the top and bottom right show this against Muslim share. Binscatters in blue show this for SC share, and those in red for Muslims. The data on public facilities comes from the Economic Census 2013 and the data on shares from the Socioeconomic Caste Census.

Figure 6

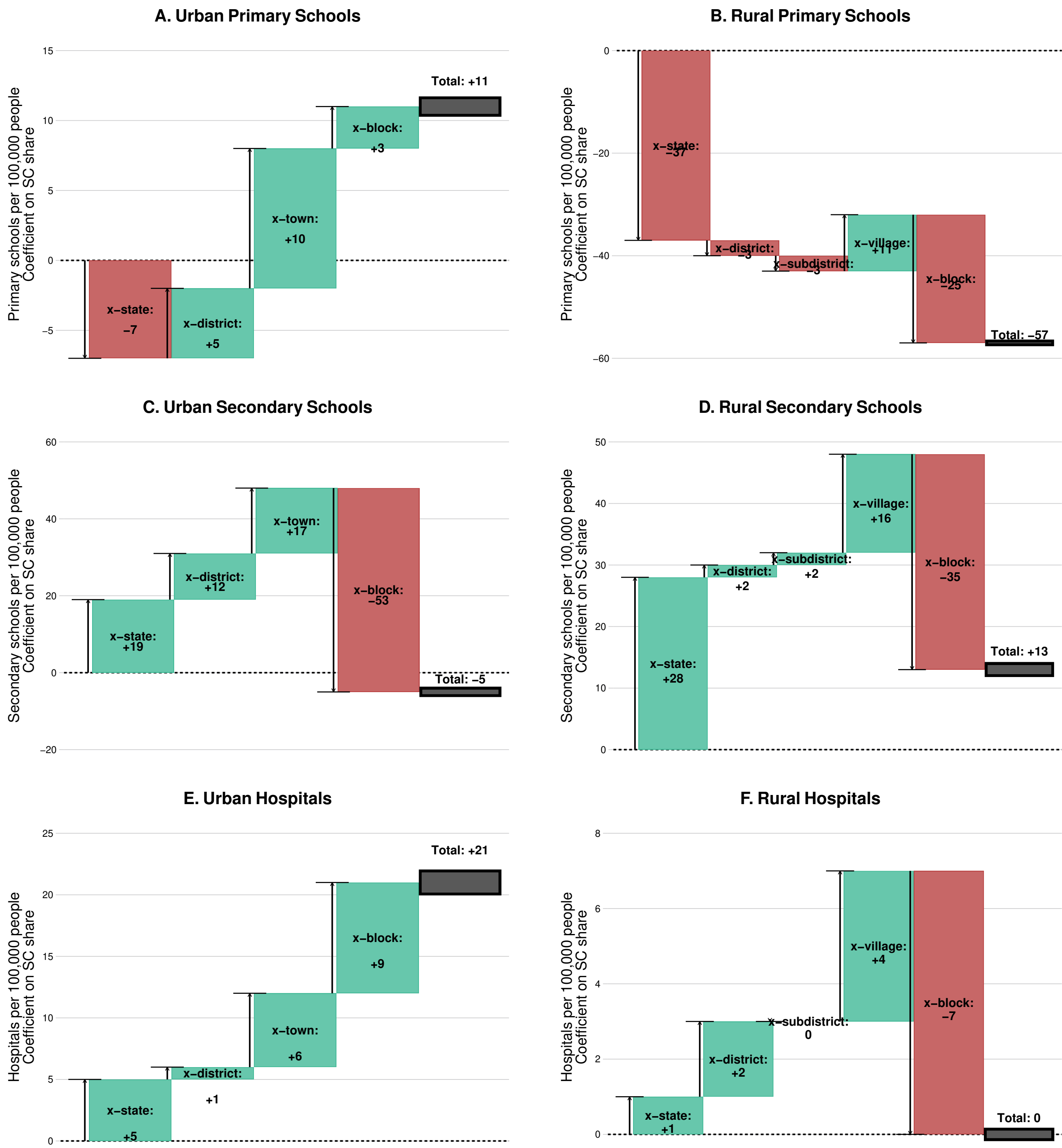
Political Economy Functions:
Disparity in Public Facilities as a Function of Muslim Share



Notes: This figure describes the cross-neighborhood relationship between a neighborhood's Muslim share and a neighborhoods access to public facilities: primary and secondary schools, and hospitals. The dark gray box shows the coefficient of a regression of a public facility indicator on the Muslim share. The boxes to the left decompose that average effect into the effect arising at the cross-state, cross-district, cross-town/village, and cross-block levels. The outcome is the number of facilities per 100,000 people. The mean of this variable in rural areas is 74 for primary schools, 15 for secondary, and 12 for hospitals. In urban areas, the means are respectively 15, 5, and 5. The data on public facilities comes from the Economic Census 2013 and the data on shares from the Socioeconomic Caste Census.

Figure 7

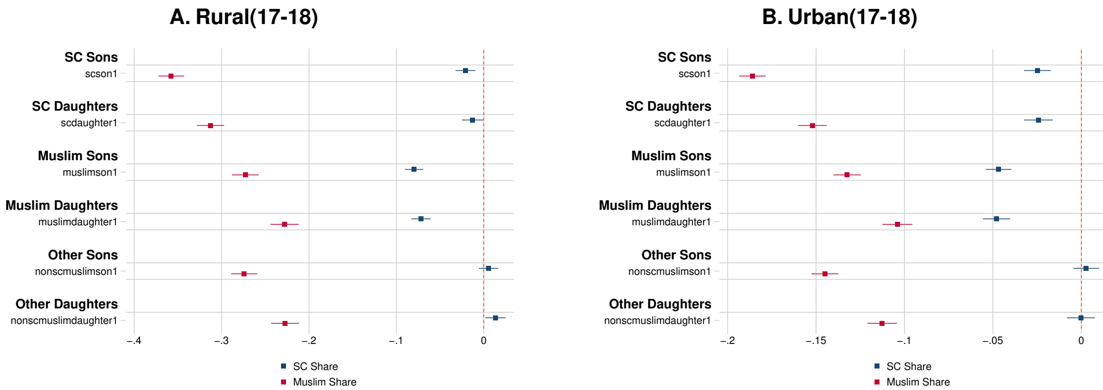
Political Economy Functions:
Disparity in Public Facilities as a Function of Scheduled Caste Share



Notes: This figure describes the cross-neighborhood relationship between a neighborhood's Muslim share and a neighborhoods access to public facilities: primary and secondary schools, and hospitals. The dark gray box shows the coefficient of a regression of a public facility indicator on the Muslim share. The boxes to the left decompose that average effect into the effect arising at the cross-state, cross-district, cross-town/village, and cross-block levels. The outcome is the number of facilities per 100,000 people. The mean of this variable in rural areas is 74 for primary schools, 15 for secondary, and 12 for hospitals. In urban areas, the means are respectively 15, 5, and 5. The data on public facilities comes from the Economic Census 2013 and the data on shares from the Socioeconomic Caste Census.

Figure 8

Educational Attainment of Children across Neighborhoods



Notes: This figure shows a coefficient plot with estimates from individual level regressions of 17-18 year old children's education (years) on neighborhood shares of SCs and Muslims, while controlling for parental education and household income/consumption and neighborhood size. The panel on the left does this for rural neighborhoods and the one on the right for urban neighborhoods. Data for the public and private facilities comes from the Economic Census 2013,while the SC and Muslim shares come from the Socioeconomic Caste Census.

Tables

Table 1
Demographics Summary Statistics Table

	Neighborhood		City/Subdistrict	
	Rural	Urban	Rural	Urban
Total Population	517.41	489.93	126423.30	57329.47
	(171.98)	(166.03)	(124796.85)	(257524.51)
Scheduled Castes Population	86	55	20774	6450
	(127)	(99)	(28417)	(27471)
Muslim Population	71	81	17506	9535
	(116)	(125)	(30651)	(52497)
Scheduled Castes (Share)	0.16	0.11	0.16	0.13
	(0.23)	(0.19)	(0.10)	(0.09)
Muslim (Share)	0.13	0.16	0.12	0.15
	(0.20)	(0.23)	(0.12)	(0.14)
Scheduled Castes (Dissim.)	NA	NA	0.63	0.60
	NA	NA	(0.13)	(0.15)
Muslim (Dissim.)	NA	NA	0.47	0.42
	NA	NA	(0.16)	(0.17)
Observations (Total)	1101321	399085	4625	3520
Observations (SC)	1101321	399085	4625	3520
Observations (Muslim)	1101321	399085	4625	3520

**Standard errors in parentheses*

Notes: This table shows summary stats for variables from the Socioeconomic Caste Census at the neighborhood-level (columns 1 and 2) and at the subdistrict/town-level (columns 3 and 4) for rural and urban areas respectively.

Table 2

Public Facilities and Consumption Statistics

	Block (Dummy)		Shrid/Subdistrict (Log Employment)	
	Rural	Urban	Rural	Urban
Public Primary Education	0.47	0.08	5.53	3.03
	(0.80)	(0.36)	(1.27)	(1.62)
Public Secondary Education	0.08	0.03	4.38	2.74
	(0.31)	(0.21)	(1.55)	(1.95)
Public Health Facility	0.07	0.03	3.42	2.33
	(0.29)	(0.25)	(1.40)	(1.89)
Consumption Per Capita(SC)	15038.87	30015.27	13091.30	21445.67
	(9247.55)	(17983.08)	(7512.10)	(7950.38)
Consumption Per Capita(Muslim)	13885.50	27036.15	12129.11	20529.86
	(8742.56)	(14702.87)	(6437.59)	(7524.82)
Consumption Per Capita(Other)	16441.03	30943.16	14812.97	25883.76
	(8177.61)	(13826.73)	(7963.96)	(8997.89)

**Standard errors in parentheses*

Notes: This table shows summary stats for public facility variables and consumption per capita from the Socioeconomic Caste Census at the neighborhood-level (columns 1 and 2) and at the subdistrict/town-level (columns 3 and 4) for rural and urban areas respectively.

Table 3

Neighborhood-level Public Facilities vs Marginalized Group Share (Urban)

	(1)	(2)	(3)	(4)	(5)	(6)
	Primary School		Secondary School		Health Facility	
	Indicator	Log Emp	Indicator	Log Emp	Indicator	Log Emp
SC Share	0.028***	0.031***	-0.006***	-0.023***	-0.004***	-0.006**
	(0.002)	(0.004)	(0.001)	(0.004)	(0.001)	(0.003)
Muslim Share	-0.004**	-0.012***	-0.011***	-0.033***	-0.010***	-0.022***
	(0.002)	(0.004)	(0.001)	(0.003)	(0.001)	(0.002)
Observations	356723	356723	356723	356723	356723	356723
Sub-district FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dependent Variable	0.066	0.117	0.023	0.063	0.022	0.043

Standard errors in parentheses

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: This table shows the regression results of public facilities on marginalized group share at the neighborhood level, for urban areas only. This is done for two different measures of public facilities: an indicator for whether or not there is the particular public facility in a neighborhood and $log(employment+1)$ in such public facilities. Columns 1 and 2 present results for primary schools, columns 3 and 4 for secondary schools, and columns 5 and 6 for health facilities. All regressions include town fixed effects and control for the log of neighborhood population.

Table 4

Neighborhood-level Public Facilities vs Marginalized Group Share (Rural)

	(1)	(2)	(3)	(4)	(5)	(6)
	Primary School		Secondary School		Health Facility	
	Indicator	Log Emp	Indicator	Log Emp	Indicator	Log Emp
SC Share	-0.006***	-0.017***	-0.007***	-0.017***	-0.002*	-0.003*
	(0.002)	(0.004)	(0.001)	(0.003)	(0.001)	(0.002)
Muslim Share	-0.086***	-0.143***	-0.021***	-0.040***	-0.014***	-0.016***
	(0.003)	(0.005)	(0.001)	(0.003)	(0.001)	(0.002)
Observations	972488	972488	972488	972488	972488	972488
Sub-district FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dependent Variable	0.328	0.535	0.069	0.145	0.058	0.079

Standard errors in parentheses

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: This table shows the regression results of public facilities on marginalized group share at the neighborhood level, for rural areas only. This is done for two different measures of public facilities: an indicator for whether or not there is the particular public facility in a neighborhood and $log(employment+1)$ in such public facilities. Columns 1 and 2 present results for primary schools, columns 3 and 4 for secondary schools, and columns 5 and 6 for health facilities. All regressions include subdistrict fixed effects and control for the log of neighborhood population.

Table 5

Individual Level Education Attainment of Children

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Rural						
	Son Ed	Son Ed	Son Ed	Daughter Ed	Daughter Ed	Daughter Ed
SC Share	-0.009	-0.006	-0.003	0.004	0.005	0.005
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
Muslim Share	-0.368***	-0.276***	-0.259***	-0.333***	-0.223***	-0.214***
	(0.009)	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)
Observations	2321240	1622446	1603088	1944447	1341761	1326036
R^2	0.16	0.28	0.30	0.18	0.32	0.34
Controls	None	Parent Ed	And Consumption	None	Parent Ed	And Consumption
Panel B: Urban						
SC Share	-0.203***	-0.034***	-0.016***	-0.203***	-0.034***	-0.016***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
Muslim Share	-0.293***	-0.137***	-0.129***	-0.263***	-0.103***	-0.100***
	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
Observations	1653931	1140195	1071443	1437992	997841	936068
R^2	0.16	0.32	0.34	0.17	0.33	0.34
Controls	None	Parent Ed	And Consumption	None	Parent Ed	And Consumption

Standard errors in parentheses

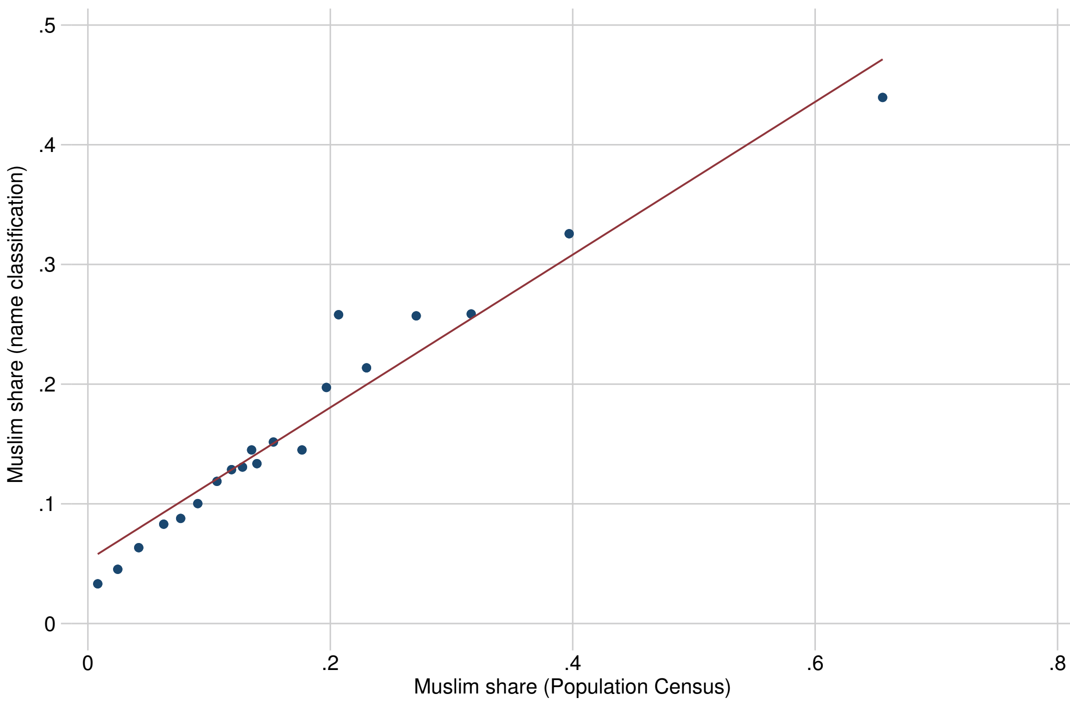
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: This table shows results from regressions on a sample of individuals taken from the SECC 2012. We use a 2 stage random sampling strategy and take a 20% sample of neighborhoods stratified at the subdistrict level and then take a simple random sample of 50% households within those neighborhoods. Education is measured in years and the dependent variable is normalised by its mean so all coefficients can be interpreted as percentages. Consumption controls are at the household level.

A Appendix A. Figures

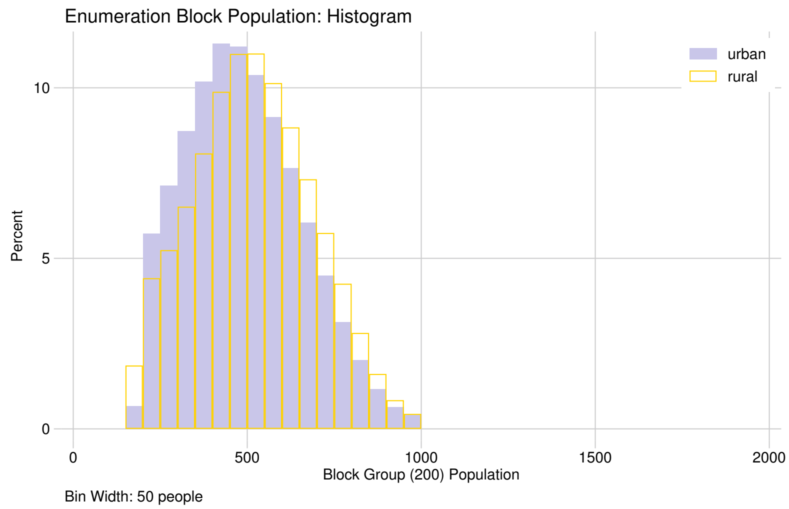
Figure A.1

Validation of Muslim Name Classification:
Subdistrict Muslim Share in SECC vs PC11



Notes: This figure is the binscatter of subdistrict-level Muslim share using our classifier of SECC names versus official 2011 Population Census Muslim share.

Figure A.2
Neighborhood Population Distributions



Notes: This figure shows the sample distribution of populations for neighborhoods (also called enumeration blocks) in urban and rural areas. Neighborhoods are excluded from the sample if they have less than 150 people or more than 1200.