

The tools for counting

As the 2011 Census approached, demands for inclusion of data on caste in Census reached a crescendo. P. Chidambaram, the Union Home Minister at the time, was opposed to collecting caste data and blocked it by claiming that it was logistically impossible for the Census, but caste information could be collected via the planned Below Poverty Line (BPL) Census, later renamed the Socio-Economic and Caste Census (SECC). Hasty inclusion of the caste question in the SECC has resulted in largely unusable data. The government tasked former NITI Aayog chairman Arvind Panagariya to look into this, but the effort has stalled. Consequently, if we want information regarding the size and characteristics of various castes in India, we must continue to look to the Census of 1931.

It is hard to imagine that the 2021 Census will not face another slew of demands for collection of caste data. It also seems likely that once again we will be unprepared for a full caste census. If we really want to collect data on caste in India and not let the discourse about Indian society be shaped by the political exigencies of colonial India, the time to plan is now.

Should we collect data on caste? Some would argue that the simple act of asking about caste creates a chasm within society. Part of this resistance comes from reaction to the preoccupation of colonial administrators-turned-arm-chair anthropologists who saw caste as the defining feature of Indian society. Colonial Censuses, beginning with the first Census in 1871, included questions about caste and used these data to divide and conquer India by first privileging Brahmins as interpreters of Indian culture and then targeting them as the roots of caste-based oppression and inequality.

G.S. Ghurye, the early 20th century pioneer of Indian sociology, reacted sharply by identifying this passion for classification as the source of anti-Brahmin movements. Veena Das, doyenne of modern Indian anthropology, also notes that the colonial Censuses via the process of recording caste generated a conception of community as a homogeneous and classifiable community and thereby influenced the processes of political representation. Consequently, post-Independence Censuses have shied away from including questions about caste.

The challenge for modern India lies in figuring out whether caste-based political mobilisation and strong sentiments for or against reservations would disappear just because we choose not to collect statistics about caste. Patels, Gujjars, Jats and Marathas do not seem to care about the lack of Census data as they demand reservations. Nor has the caste cauldron of Karnataka elections calmed because we can only roughly estimate the size of the Lingayat and Vokkaliga communities.

Our political systems, civil society and courts continue to assume that broad caste-based social categories — Dalits, Adivasis, Other Backward Classes (OBCs) and upper castes — defined largely using data from 1931 Census and a few special purpose surveys continue to shape economic conditions in 21st century India. Without accurate data at a granular level for each of these categories consisting of thousands of *jatis* (castes) and *upjatis* (subcastes), we have no way of knowing whether this is correct.

Indian society has undergone a tremendous transformation since 1931. Land ownership that bolstered the power of upper castes has lost its hold. Land fragmentation and decades of agricultural stagnation have turned many upper caste landowners into marginal farmers barely eking out a subsistence. While landlessness, once the bane of Dalit existence, has left the landless better poised to take advantage of rising rural wages, particularly construction wages. Consequently, while at a broad brushstroke the National Sample Survey (NSS) shows that mean

consumption expenditure of forward castes is higher than that of Dalits, clusters of poverty persist among forward castes. According to NSS data, the bottom fourth of forward castes are poorer than the top half of Dalits. India Human Development Survey shows that 56% of Dalit children ages 8-11 cannot read but neither can 32% of forward caste and 47% of OBC children.

Economic growth of the past century, combined with strong affirmative action undertaken by successive governments of the independent nation, may have changed relative fortunes of various groups. Some *jatis* may have managed to pull themselves out poverty and marginalisation, while others may have sunk into it. Hence, it is time to collect data that reflects the current situation.

Collection of caste data is not easy. The SECC asked interviewers to write down the name of the caste exactly as articulated by the respondent. By some reports, it has revealed as many as 46 lakh castes. Sometimes the same caste is spelt in different ways, at other times some individuals report their *jati* and others *upjati* making it difficult to create mutually exclusive categories. One cannot help but sympathise with the Ministry of Rural Development and the Ministry of Housing and Urban Poverty Alleviation which were asked to tack on a question about caste at the eleventh hour in the SECC without any preparation.

However, we have nearly three years before the Census of 2021 and are fortunate to have data from the SECC and technologies rooted in machine learning at our disposal. It would be possible to set up an expert group that uses the SECC data in conjunction with other data sources such as matrimonial advertisements and State-specific Scheduled Castes/OBC lists to make a comprehensive list of castes and condense them into meaningful categories via machine learning tools. These categories could then be validated by domain experts from the Indian Council of Social Science Research (ICSSR) institutions in various States to come up with a district specific list of castes that would cover more than 90% of individuals in any given district. Interviewers could use this precoded list to allow respondents to self-classify with a small residual group's responses being recorded verbatim and categorised later. This is very similar to the technique through which occupational and industrial classification systems are created.

Collection of data on castes is inherently risky. Politicians have long realised the advantages and disadvantages of capitalising on the sense of relative deprivation among various groups. A caste Census could easily roil the waters in ways that are hard to predict. However, once the SECC was conducted, the genie was out of the bottle. Demands are already rife for the release of these data. Conceding that these data are flawed and looking for better ways of collecting data on caste may be a way of calming the waters before the 2019 election.

It will take courage for a future government to collect data on caste and to use it to rationalise reservation policies. However, without better and more current data, our discourse on caste and affirmative action remains dominated by decisions made by the colonial administration.

Sonalde Desai is professor of sociology at University of Maryland, U.S., and Senior Fellow and Centre Director, NCAER-National Data Innovation Centre. The views expressed are personal

Sign up to receive our newsletter in your inbox every day!

Please enter a valid email address.

END

Downloaded from crackIAS.com

© **Zuccess App** by crackIAS.com