# Abstract

The political economy literature treats caste as an ontologically stable object that is entirely exogenous and not impacted by state policy or market outcomes. Such treatment of caste is consistent with how the political economy literature treats other ethnic variables. In direct contrast to the political economy literature, the sociology literature has argued for treating ethnic variables such as caste as being endogenous. Indeed, given the proliferation of the use of ethnic variables as exogenous, mediating variables like climate or topography in political economy models has resulted in calls for considered pauses, and even the need to adequately define ethnicity. The sociology literature has marshalled a range of historical examples from around the world to show why the celebrated negative relationship between ethnic diversity and political economy outcomes might be statistical artefacts from models that fail to consider the endogeneity of ethnic variables. If ethnic categories such as caste can be manipulated and constructed by state policy, then treating caste identity as an exogenous variable – especially in models that proxy state capacity to provide public goods or promote economic activity – is problematic.

This dissertation takes on this endogeneity conundrum seriously and uses a mixed-methods approach to tease out the historical political sociology of caste in the southern Indian state of Karnataka from the colonial to the present. I use a combination

of historical analysis and quantitative econometric modelling to shed new light on the endogeneity problems that undergird many of the extant political economy models. The dissertation contains three substantive essays that each look at both the causes and consequences of aggregation (or disaggregation) of caste categories across space and time.

In the first essay, I trace the evolutionary trajectories of two numerically and politically dominant castes groups of Karnataka – Lingayats and Vokkaligas. I show how the colonial caste census operations from 1871 to 1931 were central to construction of contemporary Lingayat and Vokkaliga groups. The identities of these two major castes groups in Karnataka were forged by historically contingent decisions of successive colonial Census Commissioners whose decisions on matters such as nomenclature were influenced by specific political considerations of the time. The larger narrative shows why the ahistorical treatment of ethnic groups such as caste in the political economy literature is potentially problematic.

In the second essay, I study a classical problem in urban sociology– the overlap between ethnic boundaries and spatial boundaries. I investigate if caste is a significant axis of spatial clustering in contemporary urban Karnataka. I show that caste is squarely embedded in an ethnic-geographic aggregation. I also show how the ward – the most commonly used spatial unit of analysis – is heterogeneous. Next, I develop a

framework using a new family of perfectly decomposable divergence metrics to account ethic-geographic aggregation of caste. Finally, I situate the empirical findings from this essay within the larger body of social sciences and policy literature predictions that segregation of space decreases with increasing urbanization. I find scant evidence for this commonly accepted hypothesis. I explain the empirical findings both in terms of the historical basis for the normative promise of urbanization, as well as the particularities of extant empirical models used to test the relationship between urbanization and ethnic segregation.

In the third essay, I revisit one of the central predictions of the political economy literature -- the negative relationship between social heterogeneity and economic performance. This essay makes two complementary contributions. First, I show that the relationship between heterogeneity and economic performance is likely a statistical artefact – in particular an artefact of the unit of analysis. Using a large dataset from Karnataka, I provide incontrovertible evidence for how the ethno-geographic aggregation of case is a crucial determinant of model outcomes (the negative relationship in extant literature is supported in a small minority of cases). Second, this essay develops a general framework for using ethnicity variables in a subnational setting – an area of increasing activity in recent political economy literature.