

International Journal of Population Data Science

Journal Website: www.ijpds.org



Methodological Challenges and Advances in Census-to-Census Linkage in Brazil

VInicius Prado¹ and Luiz Felipe Barros²

¹Brazilian Institute of Geography and Statistics (IBGE), Rio de Janeiro, Brazil; Harvard University, Boston, USA

²Brazilian Institute of Geography and Statistics (IBGE), Rio de Janeiro, Brazil

This study examines the methodological challenges and early advances involved in linking individual records from Brazil's 2010 and 2022 Demographic Censuses. Although Brazil has a long tradition of innovation in census operations, the linkage between consecutive census rounds remains underexplored, as observed in many developing and middle-income countries. The main goal is to identify the potential and limitations of probabilistic linkage between census datasets, particularly those related to data quality, variable availability, and structural changes across census rounds, and to document the methodological progress. Our approach comprised five stages: (1) mapping and assessing the structure, availability, and quality of identifying variables in both censuses; (2) preprocessing and harmonizing personal identifiers, household attributes, and address-related information; (3) integrating individual and household attributes to define the information hi-

erarchy for linkage; (4) designing a preliminary probabilistic linkage strategy with alternative blocking rules; and (5) conducting exploratory tests to assess feasibility and inform refinements. The Brazilian context presents specific challenges, including naming conventions, variability in information quality, and changes in household and address structures, which require tailored harmonization and linkage solutions. To assess feasibility, we implemented small-scale exploratory tests that combine territorial references with household and individual characteristics. This approach enabled us to identify practical constraints and evaluate preliminary blocking and matching rules under real data conditions. The emerging lessons and methodological insights from this work may inform the refinement of linkage strategies using data from Brazil. They may also inform data integration efforts in other developing countries facing similar structural and operational challenges.

